DOCUMENT RESUME

ED 326 239 IR 053 344

TITLE Uncovering Information: Subject Access to Information in Libraries.
INSTITUTION New York State Library, Albany.
PUB DATE 89
NOTE 72p.
PUB TYPE Collected Works - Serials (022)
JOURNAL CIT The Bookmark; v47 n3 p154-220 Spr 1989

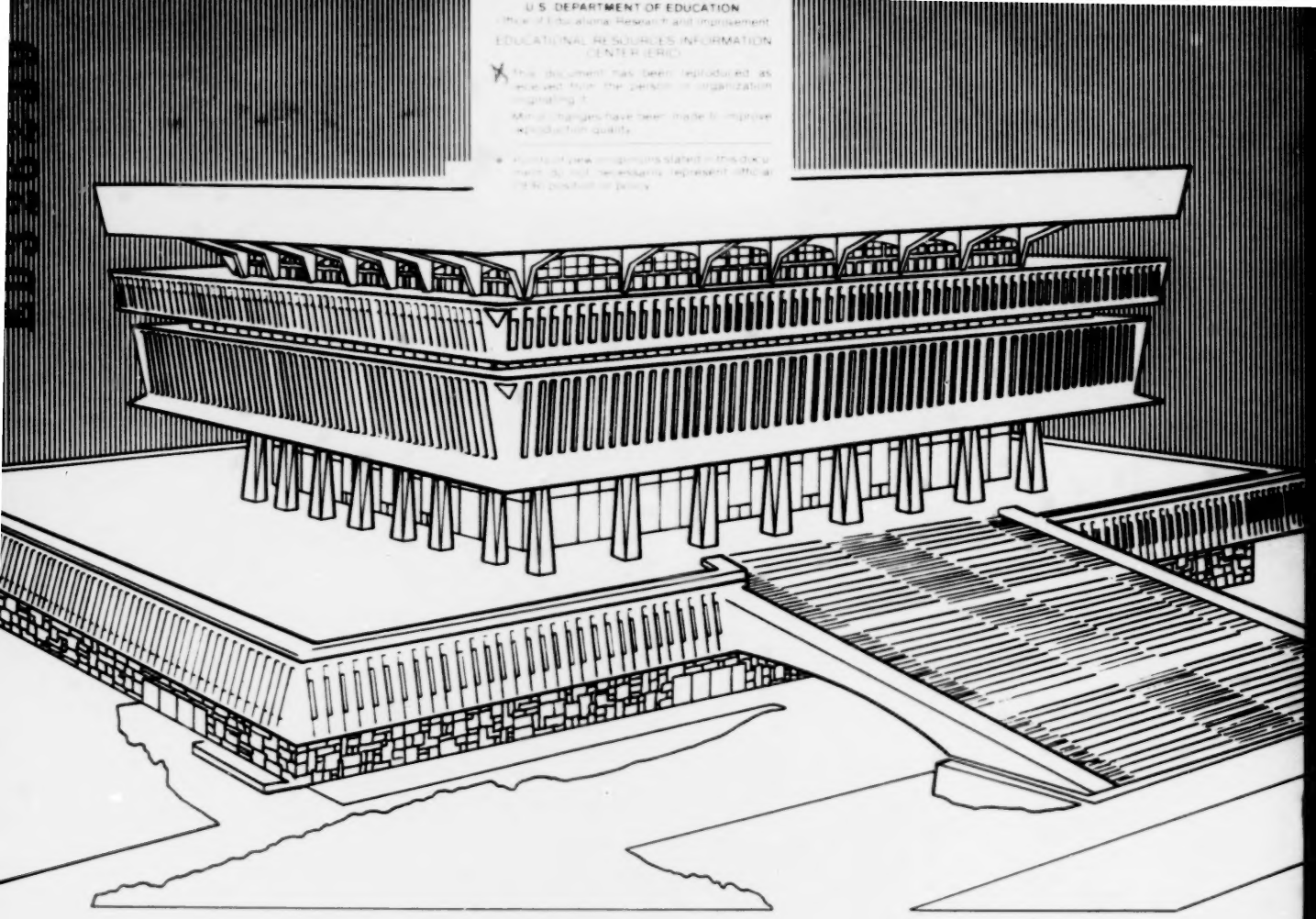
EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Academic Libraries; *Cataloging; *Classification; Elementary Secondary Education; Expert Systems; Gateway Systems; Government Publications; Higher Education; Hypermedia; *Indexing; Information Retrieval; Learning Resources Centers; *Library Automation; Library Education; Library Standards; *Man Machine Systems; *Online Catalogs; Online Searching; State Libraries; Technological Advancement
IDENTIFIERS Authority Control (Information); New York; *Subject Access to Information

ABSTRACT

The 16 articles in this theme issue focus on ways in which libraries provide subject access to their collections in a fast-changing technological environment: (1) "Uncovering Information" (Neal K. Kaske and Peter Paulson); (2) "The New York State Library's Evolving Public Access Catalog" (Alexandria Wolff and Jean Hargrave); (3) "Online Access to Subject Information at Rensselaer Polytechnic Institute" (Liz Lane); (4) "From Backwater to Mainstream: Government Publications in the Online Catalog" (Mary Redmond); (5) "Technical Standards: A Common Ground" (Walt Crawford); (6) "Hypermedia-Enhanced Subject Access to Children's Materials" (Florence E. DeHart); (7) "Interactive Authority Control" (Edward O'Neill); (8) "Dewey Into the 90s" (Peter J. Paulson); (9) "Classification and Cataloging Education" (Judith Hopkins); (10) "Classification Online and the Quest for Knowledge" (Richard S. Halsey); (11) "Appropriate User Interfaces for Subject Searching in Bibliographic Retrieval Systems" (Charles R. Hildreth); (12) "Displaying Materials in an Interdisciplinary Online World: New Approaches" (Joseph R. Matthews); (13) "Information--A Changing Resource" (John J. Regazzi); (14) "Future of Subject Searching" (Richard T. Sweeney); (15) "Expert System Interface Software in Microcomputer Workstations" (William H. Mischo); and (16) "New Challenges in Helping Students Uncover Information: The Learning Disabled Student" (Marian Traina Power, Carol Rundlett, and Myra David). Eight of the articles include references. (SD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.



The Bookmark

THE UNIVERSITY OF THE STATE OF NEW YORK / THE STATE EDUCATION DEPARTMENT / THE NEW YORK STATE LIBRARY

Uncovering Information: Subject Access to Information in Libraries

Spring 1989

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Robert M. Trombly
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

NEW YORK
THE STATE OF LEARNING

THE UNIVERSITY OF THE STATE OF NEW YORK

Regents of The University

MARTIN C. BARELL, *Chancellor*, B.A., I.A., LL.B.	Muttontown
R. CARLOS CARBALLADA, *Vice Chancellor*, B.S.	Rochester
WILLARD A. GENRICH, LL.B.	Buffalo
EMLYN I. GRIFFITH, A.B., J.D	Rome
JORGE L. BATISTA, B.A., J.D.	Bronx
LAURA BRADLEY CHODOS, B.A., M.A.	Vischer Ferry
LOUISE P. MATTEONI, B.A., M.A., Ph.D.	Bayside
J. EDWARD MEYER, B.A., LL.B.	Chappaqua
FLOYD S. LINTON, A.B., M.A., M.P.A.	Miller Place
MIMI LEVIN LIEBER, B.A., M.A.	Manhattan
SHIRLEY C. BROWN, B.A., M.A., Ph.D.	Albany
NORMA GLUCK, B.A., M.S.W.	Manhattan
JAMES W. MCCABE, SR., A.B., M.A.	Johnson City
ADELAIDE L. SANFORD, B.A., M.A., P.D.	Hollis
WALTER COOPER, B.A., Ph.D.	Rochester
GERALD J. LUSTIG, B.A., M.D.	Staten Island

President of The University and Commissioner of Education
THOMAS SOBOL

Executive Deputy Commissioner of Education
THOMAS E. SHELDON

Deputy Commissioner for Cultural Education
CAROLE F. HUXLEY

Assistant Commissioner for Libraries and State Librarian
JOSEPH F. SHUBERT

Director of the State Library
JEROME YAVARKOVSKY

Director, Division of Library Development
ROBERTA G. CADE

The State Education Department does not discriminate on the basis of age, color, religion, creed, disability, marital status, veteran status, national origin, race, or sex in the educational programs and activities which it operates. Inquiries concerning this policy of equal opportunity and affirmative action should be referred to the Department's Affirmative Action Officer, Education Building, Albany, NY 12234. Phone (518) 474-1265.

NEW YORK STATE LIBRARY **The Bookmark**

VOLUME 47 — NUMBER III SPRING 1989

UNCOVERING INFORMATION: SUBJECT ACCESS TO INFORMATION IN LIBRARIES

CONTENTS

	PAGE
Uncovering Information By Neal K. Kaske and Peter Paulson	154
The New York State Library's Evolving Public Access Catalog By Alexandria Wolff and Jean Hargrave	156
Online Access to Subject Information at Rensselaer Polytechnic Institute By Liz Lane	159
From Backwater to Mainstream: Government Documents in the Online Catalog By Mary Redmond	161
Technical Standards: A Common Ground By Walt Crawford	166
Hypermedia-Enhanced Subject Access to Children's Materials By Florence E. DeHart	168
Interactive Authority Control By Edward O'Neill	173
Dewey Into the 90s By Peter J. Paulson	176
Classification and Cataloging Education By Judith Hopkins	179
Classification Online and the Quest for Knowledge By Richard S. Halsey	183
Appropriate User Interfaces for Subject Searching in Bibliographic Retrieval Systems By Charles R. Hildreth	186
Displaying Materials in an Interdisciplinary Online World: New Approaches By Joseph R. Matthews	194
Information — A Changing Resource By John J. Regazzi	197
Future of Subject Searching By Richard T. Sweeney	201
Expert System Interface Software in Microcomputer Workstations By William H. Mischo	207
New Challenges in Helping Students Uncover Information: The Learning Disabled Student By Marian Traina Power, Carol Rundlett and Myra David	213

Joseph F. Shubert and E. J. Josey, *Editors*; Alice Britenbaker, *Editorial Assistant*.

Issued quarterly by the State Education Department, New York State Library, Albany. Sent on exchange to libraries. Individual subscriptions within the United States, Canada and Mexico are $15.00 a year; $4.00 a copy. Annual subscriptions outside of the United States, Canada and Mexico are $20.00; $8.00 a copy. Subscriptions are available from Documents/Gift and Exchange, New York State Library, Albany, New York 12230. Make checks payable to *New York State Library*.

The Bookmark is indexed in *Library Literature*, *Public Affairs Information Service* and *CALL (Current Awareness - Library Literature)*.

UNCOVERING INFORMATION: SUBJECT ACCESS TO INFORMATION IN LIBRARIES

This issue of *The Bookmark* explores how libraries are providing subject access to collections in a fast-changing technological environment. Many of these technological changes affect publishing practices, information processing and media formats and, in turn, require adjustments by librarians. Today's interdisciplinary approaches to scholarship and the increasing expectations of library patrons for additional automated services are also changing the face of subject access. This issue of *The Bookmark* presents papers that help us understand the tasks and complexities of providing better subject access in our changing environment. Authors also speculate on the future of subject access, which should aid our understanding and stimulate our thinking about how best to improve access to information and thus better serve library patrons.

The issue commences with an article on the New York State Library's efforts to provide improved subject access and the development of the Library's online public access catalog. Alexandria Wolff (Senior Librarian) and Jean Hargrave (Supervising Librarian), both staff members of the New York State Library, prepared this article. Next is an account by Liz Lane (Head, Technical Services, Folsom Library) of the actions taken at Rensselaer Polytechnic Institute to support improved subject access to their materials. In addition to online access to books and journals, Rensselaer's system (InfoTrax) provides access to a file of materials from the Institute of Electrical and Electronic Engineers (IEEE). The next paper, by Mary Redmond (Principal Librarian for Legislative and Governmental Services, New York State Library), provides a clear picture of the ways in which many technological changes are improving subject access to government documents.

The question of, and the role of, technical standards seems to always come up when one deals with subject access in a computerized setting. The role and use of standards as they relate to providing subject access is articulated clearly by Walt Crawford (Principal Analyst for Special Projects, The Research Libraries Group, Inc.). Walt has been working with and writing about technical standards that affect libraries for a number of years.

Providing subject access to children's materials has always been a real challenge, and, with the addition of hypermedia information packages to libraries, this challenge is even greater. Florence DeHart (Professor, School of Library and Information Management, Emporia State University) helps us in this area by conveying to us many of the innovations taking place in this arena.

Peter Paulson, (Executive Director, Forest Press OCLC) brings to us an interesting historical and future look at the ways the Dewey Decimal Classification system has and will be used as a subject access tool for patrons and staff. Edward O'Neill, (Senior Research Scientist, OCLC) discusses a research project that measured the variety and impact of errors in subject heading and the effects these errors have on interactive authority control systems.

Next, an overview of classification and cataloging education is presented by Judith Hopkins (Technical Services Research and Analysis Officer, University Libraries, State University of New York at Buffalo). She points out how vital a knowledge of cataloging is today and its increasing importance in dealing with automated systems. Richard Halsey, (Dean, School of Information Science and Policy, State University of New York at Albany) makes us stop and think about the role of librarian in an automated setting.

Charles Hildreth (Chief Consulting Scientist, READ Ltd.) examines the subject searching side of the online user interfaces we have today. He provides a clear and helpful distinction between querying and browsing in online systems, along with ideas on how to improve current online catalogs and information retrieval systems.

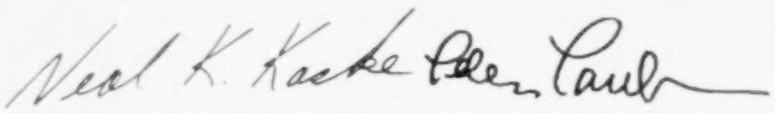
From the vantage point of a vendor, Joseph Matthews (Vice President for Marketing, Geac Computers, Inc.) gives an overview of the newer systems for displaying materials via a computer terminal. John Regazzi (President, Engineering Information, Inc.) provides us with an interesting case history of Ei's automation efforts. He also relates some useful insights into the electronic publishing field.

Richard Sweeney (Dean, Library Information Services, Polytechnic Institute of New York) gives us a wild and informative ride into the future of subject access. He does this by relating information about the basic concepts behind subject access and by reporting how the integration of expert systems, image retrieval and hypertext will change the face of subject access. William Mischo

(Engineering Librarian, Beckman Institute Library, University of Illinois at Urbana-Champaign) explains a research and development project that has created and tested an intelligent gateway software system that helps patrons who are conducting subject searches. The software is an expert system that provides access to many databases, both on and off campus. Mischo is making some of Sweeney's ideas a reality.

This issue concludes with an article authored by a team from the Academic Resource Center, C.W. Post Campus of Long Island University, Brookville. The team is composed of Marian Power (Director), Carol Rundlett (Assistant Director) and Myra David (graduate student). Their work will help all of us to become more aware of and better able to understand the world of the learning disabled student. It is clear that we are all at a disadvantage to a degree when it comes to conducting subject searches today. Methods for providing better subject searching are needed. After all, subject searching is a key route taken frequently by all to locate needed information. What we can do to improve subject access will benefit all.

Authors, THANKS! for your time, ideas and papers. Readers, we hope you will enjoy and benefit from reading this issue of *The Bookmark* as much as we have enjoyed and learned from putting it together.



Neal K. Kaske and Peter J. Paulson

The idea for this issue of this issue of *The Bookmark* grew out of a discussion with Peter J. Paulson, Executive Director of Forest Press/OCLC and former Director of the New York State Library. Mr. Paulson persuaded Neal K. Kaske, Senior Associate in the Office of Library Programs at the U.S. Department of Education to join him as guest editor. Mr. Kaske brings to this issue of *The Bookmark* a background rich in library research and automation. Neal joined the U.S. Department of Education in 1989, where he is responsible for the College Library Technology and Cooperation Grants program under Title II-D of the Higher Education Act. Before going to Washington, he worked in library education at the University of Alabama (1985-89) and at Case Western Reserve University (1983-85). He also has worked in research and research management at OCLC (1979-83) and in systems and automation at the University of California, Berkeley (1983-78).

We are also indebted to the Colorado Alliance of Research Libraries (CARL) and Executive Director Ward Shaw for permission to adapt CARL's "UnCover" logo for this issue. UnCover is CARL's article access database and service launched in September 1988. Within six months, CARL expanded UnCover to provide access to articles in more than 5,300 journals, with a database of over 180,000 article records. Librarians at Clarkson University in Potsdam were the first in New York State to secure access to CARL through NYSERNET (see page 217). Information about CARL is available from Mr. Shaw at the Colorado Alliance of Research Libraries, 777 Grant, Suite 304, Denver, CO 80203 or (303) 861-5319.

Joseph F. Shubert

THE NEW YORK STATE LIBRARY'S EVOLVING PUBLIC ACCESS CATALOG

By Alexandria Wolff, Senior Librarian, and Jean Hargrave, Supervising Librarian, Reference Services, New York State Library

The New York State Library's move to a modern facility in 1978 was the catalyst for an exciting, tumultuous decade of progressive changes in its public access catalog. These changes included:

- Photographing and microfiching 3 million cards from the 2,400-drawer catalog
- In-house design, development and implementation of an online public access catalog (OPAC)
- Retrospective conversion of 650,000 shelflist records to machine-readable format
- Upgrade of 200,000 machine records to include subject access
- Participation in Dewey decimal classification online project
- Ambitious plans for an OPAC redesign

In an incredibly short time span, the traditional on-site manual catalog was replaced by an electronic catalog, now accessible from 19 video display terminals at the State Library and statewide through nearly 100 workstations dispersed throughout New York Assembly and Senate offices, State agency libraries and worksites, New York State Interlibrary Loan (NYSILL) centers and many Capital District colleges, including Rensselaer Polytechnic Institute (RPI), Russell Sage, Siena and State University of New York at Albany (SUNYA).

CMS contains 1,070,000 bibliographic book and journal records.

Terminals at the State Library and those in Legislative offices are hard-wired to the State Education Department's UNISYS A17 mainframe. All other sites have dial-up access through UNISYS terminal emulation software known as HANDSHAKE and run on an IBM PC or its equivalent and a Hayes or Hayes compatible modem. Software, CMS access code and installation and search instructions are supplied to sites by the Legislative and Governmental Services unit of the State Library.

COLLECTION MANAGEMENT SYSTEM (CMS)

Today CMS, the State Library's electronic catalog, contains 1,070,000 bibliographic book and journal records. Full MARC records exist for monographs cataloged since 1974 through OCLC and for all serials entered via CONSER. Only author, title, publishing date, subject, added entry and call number information exists for pre-1973 cataloged materials, which number 650,000 book and journal records entered into machine-readable format through a retrospective conversion project in 1976.

Unfortunately, CMS is not "a one-stop" lookup for the State Library's immense collection of 2.1 million volumes, 4.5 million microforms and thousands of manuscripts, maps, musical scores, photographs and specialized materials. Abstracts, indexes, book catalogs, electronic databases and staff-produced tickler files supplement CMS as finding aids for locating reports, journal articles, microforms and other unclassed materials. CD-ROM databases at the newly established Electronic Reference Station enhance access to collections of Educational Resources Information Center (ERIC), National Technical Information Service (NTIS), Government Printing Office (GPO) and other major collections owned by the Library.

To enhance access to its holdings while simultaneously coping with staff limitations and budget constraints, the State Library, with the assistance of the State Education Department's Division of Electronic Data Processing, has turned to tapeloading machine-readable records onto CMS. Since 1985, SuDoc monographs have been tapeloaded monthly. Tapes of individual item cataloging were acquired for two microform sets, *Early English Books* and *Early American Medical Imprints*, and loaded onto CMS. In a cooperative cataloging venture with the University of Maryland and the University of Delaware, Library staff is currently cataloging records from the microform collection, *Goldsmiths'-Kress Library of*

Economic Literature, using OCLC. Upon completion of this project, the Library will receive a tape of all the cataloged items to load onto CMS.

The Library's OPAC offers more powerful searching than was possible with our traditional card catalog. In addition to exact known author, title, author/title and subject searching, records in the CMS bibliographic database may be accessed by subject/title key words and by call number. Author key word access is currently being developed and should be available in early 1990. Function keys are used to select the type of access desired.

Search keys are used to locate records for known author, title or author/title. The key consists of portions of a name or title; for instance, MICH,JAM, is the author key for James Michener. The computer knows which type of search to perform by the number of commas in the search key. Users have a choice of menu or command driven mode for subject and title key word searching. Call number searching may only be performed through commands.

User friendly menu screens list all appropriate actions at any point in the search. Additional help is available upon request. Specific error messages are automatically displayed to assist users when they err. In the menu mode, key words are automatically combined; the computer looks for all word or word combination within the same heading or title. Words may be truncated to retrieve all variation in word endings. A search by exact Library of Congress subject heading is also possible. This method requires the user to first consult the Library's printed list of subject headings and to select the six digit number associated with the heading or headings desired. The six digit number is used in lieu of the actual words of the heading.

The advanced command mode offers staff and frequent library users more powerful, precise and streamlined search capability. In this mode, staff and patrons can use Boolean operators (AND, OR, NO, ADJ), nesting and word truncation and can specify that subjects or titles be limited to specific word combinations.

Key word searching has enhanced subject access: Patrons need not know an exact subject or title because the computer looks for the search term or terms anywhere within the subject or title field. The ability to subject search CMS on words from either the assigned subject heading or title also enhances retrieval as natural language terminology not reflected in Library of Congress controlled vocabulary may be present in the title. Shelflist "subject" browsing is now possible through the call number search. This is extremely valuable to our patrons as most of the Library's collection is housed in closed stack areas.

Subject searching on CMS is limited by the lack of an online thesaurus of Library of Congress subject headings and online authority control. Subject searching is also impeded by the limited number of subject headings per record and the imprecision of those headings. Initiatives to enrich subject access to monographic records, such as the assignment of subject descriptors from a work's table of contents or index or through Dewey decimal classification schedule online, can only be achieved by cooperative efforts. Improving access to the Library's OPAC and collection is the challenge and potential reward of our next decade.

THE FUTURE: THE NEW YORK STATE INFORMATION NETWORK (NYSIN)

Recently, reference librarians, catalogers, collection management staff and computer specialists pooled their expertise to create a proposal for an integrated, multi-institution system to replace the Library's OPAC. Planning involved not only an expert knowledge of library operations and technical capabilities but also an understanding of the Library's long-term goals and objectives and consideration of the interface between users and the system.

The trend toward searching and using library resources off-site is significant for the New York State Library because of its concern for service to residents statewide and to a State government workforce that is deployed throughout the State. The option of searching the State Library's holdings from State government offices and from library systems' interlibrary loan desks has facilitated the use of interlibrary loan and strengthened the Library's ability to extend its resources to researchers far beyond the confines of the Cultural Education Center in Albany.

NYSIN would help State agency libraries and the State Library achieve a common goal: improving access to information resources to meet the needs of State government.

PROPOSAL FOR MULTI-INSTITUTION DEVELOPMENT

As a replacement for the current online catalog, staff have envisioned an information system based on

the collective information resources of New York State government libraries and information centers as well as the State Library's holdings. The name suggested for this new system conveys its expanded scope: the New York State Information Network (NYSIN). NYSIN would help State agency libraries and the State Library achieve a common goal: improving access to information resources to meet the needs of State government. The shared system would also encourage resource sharing and cooperative collection development to promote the most economical use of materials.

Using NYSIN, researchers would be able to access multiple information resources in a uniform way using one automated system. Government researchers at office workstations in State agencies and the Legislature would not only have improved access to materials in their own agencies, but would be able to view them in the context of a broader information environment.

The mix of resources from a variety of institutions would provide an added dimension for users. For instance, the proposed government information system would include the holdings of the State Library, the State Archives and Records Administration (SARA), the State Museum and some other libraries and information centers in New York State government. In addition to library materials, the system would inform patrons about the existence of archival records and collections of artifacts, specimens and materials. Presenting information on complementary holdings of these institutions would underscore the relationship of different types of resources and reveal new vistas for researchers.

INTEGRATION OF SYSTEMS

NYSIN would also bring together crucial information from different automated systems used at the State Library for circulation, acquisitions, serials control, cataloging, authority control, inventory, binding and binding notification. Patrons would be informed immediately about the general availability of an item including circulation, acquisition and binding status. Off-site users could route requests automatically to the State Library. The Library could send materials and data directly to patrons via facsimile transmission and electronic mail.

ENHANCED SUBJECT SEARCHING

Boolean logic would be used to search the system. Using a set of full or abbreviated commands recommended by the National Information Standards Organization (NISO), the system would provide an online tutorial, a menu of help screens and on-screen prompts to assist the user. Many features, such as the ability to combine different types of searches in one search statement and limit searches by date or date range, language, type of material and location code were chosen to enhance the successful interface between the user and the system.

Two major search types would be available and would be performed using full or abbreviated versions of NISO commands, FIND and BROWSE. The FIND command isolates a results set that meets the search criteria. A BROWSE command reveals a sorted list of records in call number order or an alphabetical listing of subject headings through which the user can search forward or backward. Subjects could be searched using key word search for words in any order linked by Boolean operators in titles, note fields or subject headings. Subjects could also be searched using phrase searching for words in exact order in titles and subject headings.

FUTURE FORECAST

Whether or not the New York State Information Network is developed as proposed depends on a variety of circumstances. Undeniably, its implementation would be a major undertaking with significant staff and budgetary implications. New and evolving technologies may suggest other directions for automation development. It is clear, however, that certain features of the current proposal are fundamental to the mission of the State Library. Based on that premise, it can be predicted that future development will include sharing resources of other institutions in State government, integrating automated systems in use at the State Library, extending statewide access to the system and enhancing subject access and other features that will promote patrons' successful, independent use of the system.

ONLINE ACCESS TO SUBJECT INFORMATION AT RENSSELAER POLYTECHNIC INSTITUTE

By Liz Lane, Head, Technical Services, Folsom Library, Rensselaer Polytechnic Institute, Troy, NY

Rensselaer Polytechnic Institute, located in Troy, NY, is an institution of higher education offering undergraduate and graduate degrees in science, engineering, business and the social sciences. Currently the undergraduate enrollment is 4,200 and the graduate enrollment is 1,500. There are two libraries on the campus: Folsom Library, which houses the main collection of over 370,000 volumes, and the Architecture Library, which contains specialized materials in art and architecture.

In the late 1970s, after the completion of a new library building, the library explored the possibilities for automating library operations. At that time the library administration was authorized to plan for library automation and proceeded to form an automation committee, hire a systems analyst and draft a plan for a locally developed automated library system.

A serials file, designed as part of a student project, served as a test of the SPIRES software and became the first component of the system. The file is accessible by any word in the title and by general subject terms. Later the Library Automation Committee designed an acquisitions component that provides an order file, a vendor file and a fund file. The order file is searchable by author, title and publisher, but not by subject.

InfoTrax, Rensselaer's automated information system, became available to the public in 1984. InfoTrax incorporates the earlier serials and acquisitions components. The system, designed and implemented by library personnel, gives access to the libraries' collections of books, journals, scores, sound recording and slides, as well as the Institute of Electrical and Electronic Engineers (IEEE) file, a homework file, news, message, an automated circulation system and an information request service. Soon computer software and manuscripts will be added. At a later date the remaining formats of audiovisual materials and maps will be incorporated. All existing bibliographic files except IEEE, which is not in MARC format, will be merged for a new release of the system in the fall 1989. The system runs on the campus IBM 3190 mainframe computer and is accessible from terminals both on and off campus. Patrons can search by traditional author, title and subject approaches as well as by report numbers, architects and composers. Full Boolean searching is available.

Of particular interest is the IEEE file that lists the engineering society's publications. This file includes the citations and abstracts of IEEE journals and conference publications and can be searched by author, title (both journal and conference title) and subject.

The results of the Council on Library Resources' Online Project's studies on the use of online catalogs became available at the time the InfoTrax system was being developed. Consideration of the findings from the studies influenced the design of InfoTrax. Thus, the catalog, journals, slides and IEEE files are searchable by subject in InfoTrax.

For most of the files, the *Library of Congress Subject Headings* (LCSH) is the source of subject headings. In the slides file, subject headings are assigned using the *Art and Architecture Thesaurus* (AAT); if an appropriate subject heading is not in AAT, one is chosen from LCSH. Use of these standardized lists of terms provides structure and maintains consistency in subject headings to improve subject access.

InfoTrax offers search strategies that facilitate subject retrieval.

Studies indicate that many library users are not aware that libraries use a standard list of subject headings. This finding has implications for the way that library users construct their searches. Also, searchers tend to enter broader terms whereas established cataloging practices advocate assigning subject headings as specifically as possible. The possibility exists that searchers, unaware of the use of a controlled vocabulary, will not retrieve appropriate results because they do not enter the terminology used within the system. Both of these situations contribute to unsuccessful subject searching.

To expand the subject access capability and ameliorate the limitations inherent in a structured subject vocabulary, InfoTrax offers a number of other search strategies that facilitate subject retrieval. Just as in the traditional card catalog where an individual could search other subjects listed on the catalog card, a user can search these same terms and combinations of terms online much faster and do so with more precision. Searchers may use the Boolean operators of "and," "or" and "and not" within a search strategy to refine a search. An "about" index combines terms available from both the subject index, containing LCSH terms, and the title index, containing significant words from the title. The "about" search functions like key word searching in other online catalogs. Also, library users can search by call number to retrieve related materials in specific subject classifications. This is particularly helpful to identify other library materials within the same subject areas. If a library user cannot think of a term that describes subject need, there is a "SCAN" command to view a list of terms or numbers within a specific index that can help to pinpoint a term. All of these search strategies were designed to provide alternative methods of accessing library materials.

With a locally developed system and programmer on the library staff, adjustments and changes can be incorporated in a short time without dealing with vendor bureaucracy. Since its introduction in 1984, InfoTrax has undergone some major and minor revisions. It is viewed as an evolving system rather than an as an established system. Removal of the card catalogs in 1988 significantly increased the number of reference questions on how to use the online system. It seems that patrons had continued to consult the card catalog in spite of the notice that no new cards had been filed into it for almost three years. Increased use requires that the system be as helpful as possible and enhancements are planned to improve the usefulness of the system responses. Plans are also being made to expand the system to include other campus files. Some of the files to be added are the student and faculty/staff directories, the RPI undergraduate and graduate college catalogs, course offerings and class schedules.

One feature that InfoTrax lacks is authority control. Such a feature would facilitate the maintenance of consistency in names, subjects and series. The majority of Rensselaer's retrospective conversion was completed from 1980-85 between two major AACR 2 conversions by OCLC. Therefore, there are hundreds of bibliographic records in the online file that may not be retrieved unless a searcher uses more than one search strategy to retrieve alternate or earlier name or subject terms. The implementation of authority control will provide cross references and improved subject access to the RPI libraries' collections of materials.

FROM BACKWATER TO MAINSTREAM: GOVERNMENT DOCUMENTS IN THE ONLINE CATALOG

By Mary Redmond, Principal Librarian, Legislative and Governmental Services, New York State Library

Publications of the United States government comprise an unparalleled corpus of diversified and valuable information. Using the thousands of titles produced each year, readers can do everything from following the progress of a piece of legislation or locating sites of toxic chemical emissions, to planning their summer vacations or determining the nutritional content of their favorite foods. Government officials, researchers, students and consumers are among those who benefit from the information contained in Federal documents.

Libraries of all types and sizes routinely collect U.S. government publications to serve their clients. The scope and usefulness of their contents, coupled with their usually relatively modest purchase prices, make Federal documents an informational and financial bargain for the cost-conscious library.

The most complete and systematic collections of Federal documents are located in the U.S. Government Printing Office's network of almost 1,400 depository libraries nationwide. Under provisions of Title 44 of the *U.S. Code*, these libraries receive Federal documents without charge. In exchange, the libraries agree to process and house the publications and to make them available for the free use of the general public. Academic, public, law, court, Federal and state libraries participate in the Government Printing Office depository program.

About 1,350 of the libraries are selective depositories. These institutions choose the material they wish to receive from a list of over 7,000 item numbers arranged by issuing agency. Selective depositories must retain depository publications, except for ephemeral or superseded material, for a minimum of five years.

An additional 53 libraries serve as regional depositories. These libraries must collect and permanently retain all publications (except ephemera and superseded material) offered under the depository program. In addition, regionals serve selective depositories in their area (usually state) with backup interlibrary loan, reference and consultation services. They also assist selectives with the disposal of depository material over five years old.

The GPO Depository Library Program is an efficient mechanism for disseminating information. From October 1, 1987 through August 31, 1988, the Government Printing Office distributed 26,191 titles to depository libraries. Impressive as this figure is, it represents a decline from 39,490 titles for the comparable period in 1986/87.[1] Since 1981, the amount of material supplied to depositories has declined because of cutbacks in government publishing and problems with microfiche quality control.[2] Depending on their item selections, depositories can add thousands of pieces to their collections each year.

As an aid in arranging collections of government publications, the Superintendent of Documents has created a system for classifying documents by issuing agency. The so-called SuDocs classification system provides a ready-made method for libraries to shelve and retrieve their Federal documents materials.

The Superintendent of Documents classification number is one of the identifying elements included in the *Monthly Catalog of United States Government Publications* citations. Preparation of this title, which also contains standard bibliographic information for Federal documents, is required by Title 44, Section 1711 of the *U.S. Code*.

Government documents have not been exploited to their full potential.

Despite the existence of the GPO Depository Library Program and accompanying finding aids, government documents have not been exploited to their full potential. Federal publications traditionally are not consulted or circulated on a level with other portions of library collections. For example, the selective depository library at the College of Wooster (OH) discovered that only 0.5 percent of the government documents collection was circulated each year compared to 18 percent of the Library of Congress-classified books and periodicals. This same library reported that government documents accounted for only 2.5

percent of library materials circulated from July 1, 1985 through June 30, 1986, although they represented 46 percent of the material available for circulation.[3] These experiences are not unique.

The lack of use is not surprising. In the words of one author, "cataloging of government documents in most libraries seems to have been something of a stepchild to the cataloging of more traditional library materials."[4] The 1985 Biennial Survey of Depository Libraries found that only 19.1 percent of depositories cataloged three quarters or more of their Federal documents. Instead, most libraries file their government documents in SuDocs order and rely on the *Monthly Catalog* and other indexes for access.

For all but the most knowledgeable and sophisticated of patrons, this approach raises formidable obstacles to access. As researchers at the University of Illinois have observed, "When either all depository documents or only those defined as 'low use' are housed in a separate area, users must be made aware of the existence of this separate collection and of whatever special files the library maintains to provide access to it. Users must also be introduced both to the government-produced indexes and bibliographies that are keys to the collection and to specialized indexes produced by commercial publishers. Furthermore, they must come to understand (presumably, with the help of librarians) the complexity and changing character of government organization and the intricacies of documents bibliography. The task of user education is a formidable challenge that few documents libraries or reference departments are fully equipped to meet."[5]

Fortunately, more libraries are now beginning to include government documents in their catalogs. Developments in recent years have opened up possibilities for libraries with online catalogs. The major methods are 1) direct loading of *Monthly Catalog* tapes, 2) contracting with a vendor to use *Monthly Catalog* information, and 3) individual (piece by piece) cataloging on OCLC or other bibliographic utilities. The February 1990 anticipated implementation of OCLC's GOVDOC, a new service for bibliographic control and resource sharing of Federal depository materials, will add yet another new option for libraries. (No evaluation of this service has been made as of this writing.)

MONTHLY CATALOG TAPELOADING

Long a mainstay of bibliographic access to Federal documents, the *Monthly Catalog of United States Government Publications* is now created by GPO by means of OCLC. The tapes used to produce the *Monthly Catalog*, beginning with the July 1976 issue, are for sale by subscription from the Library of Congress Cataloging Distribution Service. Availability of these tapes, GPO adoption of Anglo-American Cataloging Rules and MARC formats, and cooperation with the Library of Congress have greatly improved access to bibliographic information for Federal documents. The New York State Library has been loading *Monthly Catalog* tapes into its online catalog since 1985, thus facilitating access to the largest collection of Federal documents in the State. Similarly, the Suffolk Cooperative Library System in Bellport, NY has incorporated information from *Monthly Catalog* tapes and its depository selections into a CD-ROM catalog of its collection.

Monthly Catalog *tapes are not the answer to all access problems.*

This does not mean, however, that use of *Monthly Catalog* tapes is the answer to all access problems. The literature of government documents librarianship is replete with examples of the pitfalls and problems encountered in using *Monthly Catalog* tapes. These include the existence of separate (rather than collective) records for each issue of a serial or part of a monographic set, poor documentation, multiple use of the same OCLC number and time lag between receipt of an item and availability of cataloging information. In the *Administrative Notes* of August 21, 1989, GPO announced plans to reformat *Monthly Catalog* tapes to address some of the issues raised by current and potential users.[6] For now, however, the complexities of using *Monthly Catalog* tapes are such that only libraries with access to considerable automation expertise can realistically expect to use the tapes successfully.

VENDOR CONTRACTING

For libraries without the necessary in-house expertise to deal with *Monthly Catalog* tapes, contracting with vendors offers an efficient and cost-effective way of adding government document bibliographic information to online or traditional catalogs. MARCIVE (San Antonio, TX) and Brodart are examples of companies offering these services. Vendors work with a library to specify the item numbers in its depository profile, then use the cataloging information from *Monthly Catalog* tapes to produce a tape that can be loaded into the library's online catalog or used to generate cards or COM. Likewise,

nondepository libraries can also avail themselves of these services for ongoing and retrospective conversion projects.

BIBLIOGRAPHIC UTILITY CATALOGING

A third option is individual cataloging of documents by a library on a bibliographic utility such as OCLC. The archival tape is then used to generate an online or other type of catalog.

This method of processing, like the others, has both positive and negative aspects. On the one hand, there is the advantage of timely access (one study found records for 75 percent of depository documents within one month of the date of receipt.)[7] Also, entry of holdings into an online bibliographic utility provides a nationwide record of location for interlibrary loan purposes.

On the other hand, this system requires more financial and staffing commitment than other methods of online catalog access. Libraries will have to do careful analyses to determine the method that matches their needs and their pocketbook.

RESULTS

When government documents are integrated into the main catalog, the results are dramatic. Trinity University in San Antonio, Texas, for example, reported a 300 percent increase in documents usage once records were included in the circulation system.[8] Likewise, the Suffolk Cooperative Library System has noted a threefold growth in documents statistics since the inclusion of Federal publications in its CD-ROM catalog.[9] At the State Library of Pennsylvania, use of the *Monthly Catalog* has lessened significantly. "Staff and patrons alike will wait for a terminal rather than begin a search in the *Monthly Catalog*."[10]

ALTERNATIVE ONLINE ACCESS

While inclusion of Federal document records in a library's online catalog is the most direct way to improve access to its government publications holdings, other online services offer alternative methods for locating these materials. Vendors like DIALOG and BRS produce online equivalents of the *Monthly Catalog* that can be searched by subject, issuing agency, title, report number, GPO stock number and other variables. The sophistication and flexibility of the search software make these services powerful research tools.

As one researcher points out, "This enhanced access capability must be weighed against several factors: (1) the holdings of the library are not reflected, (2) the service is expensive with at least part of the cost passed on to the user, (3) searches must be performed by a staff member, and (4) the Superintendent of Documents classification number is printed without colons and/or slashes...; and, most importantly, bibliographic records for U.S. documents are still not in the catalog, which should ideally reflect all material in a library. However, as search systems become more user friendly and do not require the intervention of a skilled searcher, and as libraries subsidize online access, the use of the GPO tapes through these vendors presents the library with an enhancement to or a temporary substitute for access via the online catalog."[11]

Several vendors have begun offering Monthly Catalog *access in CD-ROM format.*

Several vendors have begun offering *Monthly Catalog* access in CD-ROM format. These products offer advantages of end-user searching and standard SuDocs number display not available in online commercial versions. For a review of some of these products, see *Library Hi Tech*, n63, September 1989, pp. 1-5. Bibliographic utilities like OCLC and RLIN, both of which include *Monthly Catalog* tapes back to 1976, offer yet another way to improve online access to Federal publications. In addition to facilitating the cataloging of government documents, bibliographic utilities are also powerful reference tools for public services staff.

Public service staff with access to an RLIN terminal can search by indexes including name, title phrase, title word, corporate/conference name phrase, corporate/conference name word, or Library of Congress subject phrase. With added abilities of Boolean searching, truncation and limitation to GPO records, RLIN users can locate both broadly based and precise citations to Federal documents. At the Colorado State University Libraries, RLIN searches have supplanted both printed and other online database tools as the preferred method for locating Federal documents.[12]

Although it does not offer all of the search capabilities of RLIN, OCLC similarly can be consulted for verification of Federal publications. Among access points are issuing agency, title, personal author, author/title and SuDocs number. As OCLC moves toward broadening the search capabilities of its database, it will become an even more powerful tool for locating government documents.

CONCLUSION

The online revolution has had, and continues to have, profound effects on access to all types of information in libraries. Not the least of its effects is the opportunity to enhance access to publications produced by the Federal government. As more and more libraries avail themselves of an array of possibilities that were unavailable even a few years ago, the cherished principle of public access to government information will move closer to becoming a universal reality.

REFERENCES

[1] *Administrative Notes*, v8, n18, October 1987. p. 11.

[2] *Administrative Notes*, v9, n8, October 1988. p. 5.

[3] Powell, Margaret S., Johnston, Deborah Smith, and Conrad, Ellen P., "The Use of OCLC for Cataloging U.S. Government Publications: A Feasibility Study," *Government Publications Review*, v.14, 1987. pp. 61-76.

[4] Plaunt, James R., "Cataloging Options for U.S. Government Printing Office Documents," *Government Publications Review*, v12, 1985. pp. 449-456.

[5] Watson, Paula D., and Heim, Kathleen M., "Patterns of Access and Circulation in a Depository Document Collection Under Full Bibliographic Control," *Government Publications Review*, v11, 1984. pp. 269-292.

[6] See:

Bower, Cynthia E., "OCLC Records for Federal Depository Documents: A Preliminary Investigation," *Government Information Quarterly*, v1, 1984. pp. 379-400.

Cornwell, Gary, "GPO Cataloging Records: Background and Issues," *Documents to the People*, v17, 1989. pp. 83-85.

Mooney, Margaret T., "GPO Cataloging: Is It a Viable Current Access Tool for U.S. Documents?" *Government Publications Review*, v16, 1989. pp. 259-270

Myers, Judy E., "The Government Printing Office Cataloging Records: Opportunities and Problems," *Government Information Quarterly*, v2, 1985. pp. 187-192.

[7] Powell, et. al., *Op. Cit.* p. 70.

[8] U.S. Congress, Office of Technology Assessment, *Informing the Nation: Federal Information Dissemination in an Electronic Age*, Washington, DC, U.S. Government Printing Office, 1988. p. 146.

[9] Eidelman, Diane, Suffolk Cooperative Library System, Bellport, NY, personal communication, May 1989.

[10] Ingraham, Alice, "Making the Monthly Catalog Obsolete: Government Documents Reference in an Online Catalog Environment," *Documents to the People*, v15, 1987. pp. 228-231.

[11] Swanbeck, Jan, "Federal Documents in the Online Catalog: Problems, Options, and the Future," *Government Information Quarterly*, v2, 1985. pp. 187-192.

[12] See Ernest, Douglas J., "Accessing Federal Government Publications with RLIN," *Government Publications Review*, v15, 1988. pp. 237-244.

BIBLIOGRAPHY

Bower, Cynthia E., "OCLC Records for Federal Depository Documents: A Preliminary Investigation," *Government Information Quarterly*, v1, n4, 1984. pp.379-400.

Cornwell, Gary, "GPO Cataloging Records: Background and Issues," *Documents to the People*, v17, n2, June 1989. pp.83-85.

Ernest, Douglas J., "Accessing Federal Government Publications with RLIN," *Government Publications Review*, v15, n3, May/June 1988. pp.237-244.

Garner, Diane, and Smith, Diane, "Government Databases: A Sampler," *Government Publications Review*, v12, n2, March/April 1985. pp. 143-154.

Hardgrove, David J., ed., "Automatic Bibliographic Control of Government Documents: Current Developments," *Documents to the People*, v15, n4, December 1987. pp.224-246.

Hernon, Peter, McClure, Charles R., and Purcell, Gary R., *GPO's Depository Library Program: A Descriptive Analysis*, Norwood, NJ, ABLEX, 1985.

Hernon, Peter, and McClure, Charles R., *Public Access to Government Information: Issues, Trends, and Strategies*, 2d ed., Norwood, NJ, ABLEX, 1988.

Jamison, Carolyn C., "Loading the GPO Tapes — What Does It Really Mean?" *Government Publications Review*, v13, n5, September-October 1986. pp. 549-559.

Mooney, Margaret T., "GPO Cataloging: Is It a Viable Current Access Tool for U.S. Documents?" *Government Publications Review*, v16, n3, May/June 1989. pp. 259-270.

Myers, Judy E., "The Government Printing Office

Cataloging Records: Opportunities and Problems," *Government Information Quarterly*, v2, n2, 1985. pp.187-192.

Plaunt, James R., "Cataloging Options for U.S. Government Printing Office Documents," *Government Publications Review*, v12, n5, September/October 1985. pp. 449-456.

Powell, Margaret S., Johnston, Deborah Smith, and Conrad, Ellen P., "The Use of OCLC for Cataloging U.S. Government Publications: A Feasibility Study," *Government Publications Review*, v13, n1, 1987. pp.61-76.

Stanfield, Karen, "Documents Online: Cataloging Federal Depository Materials at the University of Illinois," *Illinois Libraries*, v68, n5, May 1986. pp.325-329.

Swanbeck, Jan, "Federal Documents in the Online Catalog: Problems, Options, and the Future," *Government Information Quarterly*, v2, n2, 1985. pp.187-192.

Walbridge, Sharon, "OCLC and Improved Access to Government Documents," *Illinois Libraries*, v68, n5, May 1986. pp.329-332.

Watson, Paula D., and Heim, Kathleen M., "Patterns of Access and Circulation in a Depository Document Collection Under Full Bibliographic Control," *Government Publications Review*, v11, n4, July-August 1984. pp.269-292.

THE STATE LIBRARY AGENCIES: A SURVEY PROJECT REPORT, 1987

The new biennial compilation *The State Library Agencies: A Survey Project Report, 1987*, compiled and edited by Amy Kellerstrass at the Illinois State Library in cooperation with the Chief Officers of State Library Agencies, is now available. Arranged alphabetically by state, information is provided on administrative location in state government, powers and duties, major functional areas, notable activities, personnel and publications. An organization chart for each state agency is also presented. Appendices include significant comparative data for each state on functions and services, materials collections, operating receipts and expenditures, staffing and amount and use of Federal and state aid to libraries.

Requests for copies of this publication are $22.50 for ASCLA members and $25.00 for others. Order from: ASCLA Publications, American Library Association, 50 E. Huron St., Chicago, IL 60611.

TECHNICAL STANDARDS: A COMMON GROUND

By Walt Crawford, Principal Analyst for Special Projects, The Research Libraries Group, Inc.

When you work to provide subject access in online systems, you rely on any number of technical standards, although you may never have reason to concern yourself with them. On one hand, technical standards have little or no direct relevance for subject access *per se*. On the other hand, technical standards make online systems possible and may, in the future, make subject access somewhat more straightforward.

BACKGROUND

Technical standards define and, by defining, serve to improve communication and simplify development. Widespread use of common definitions saves money, time and trouble; that's the primary payoff for developing technical standards.

To take a simple example, when you purchase a personal computer or hair dryer for use in the United States, you need not spend hours (or days) negotiating to assure that the unit will accept a voltage that is available to you and that it has a connector you can use. Instead, you take it for granted that it will have a standard U.S. three-prong plug — or two-prong plug that fits into a three-hole outlet — and will run on 100 to 120 volts of alternating current at a nominal frequency of 60 Hz. That assumption takes into account hundreds of technical standards developed over the years, some of them reflected in laws and regulations.

Every online system depends on hundreds or thousands of technical standards.

Every online system depends on hundreds or thousands of technical standards — from those defining raw materials and electronic components, through many safety standards, to standards for data storage systems, character definitions and telecommunications. You don't need to know what all these standards are; few of them have any obvious relevance to the library field, but you would surely miss them if they weren't there.

RELEVANT STANDARDS

Traditional subject access relies on the *Library of Congress Subject Headings* (LCSH), *Medical Subject Headings* (MeSH) and other subject thesauri or pseudo-thesauri and on Dewey Decimal Classification, Library of Congress Call Numbers and other classification schemes. These are all codified forms of common practice, but not technical standards.

The technical standard that has done most to encourage online library systems and, thus, online subject access is ANSI Z39.2, Bibliographic Information Interchange. That standard forms the basis for the USMARC bibliographic format and authority format, without which we would not have the current wealth of machine-readable bibliographic and subject information.

An extension of USMARC to carry classification schedule information (currently being designed by the Library of Congress) will further the role of USMARC as a direct basis for subject access by providing a common communications format to distribute DDC and LCC schedules in machine-readable form.

Standardization is unwise while there is clear room for major improvement through innovation.

ANSI Z39.2 is maintained by the National Information Standards Organization (NISO), the accredited organization that develops standards in the fields of libraries, publishing and information science. Other NISO standards that can influence subject access, to the extent that they are used, include Z39.4, Basic Criteria for Indexes and Z39.19 — Guidelines for Thesaurus Structure, Construction, and Use.

Many other Z39 standards affect online access in general, but not in ways directly relevant to subject access. Each standard provides a common understanding, a set of shared explicit assumptions on which systems can be built.

Some have called for technical standards for basic displays in online catalogs. Arguments can be made for and against such standards on theoretical grounds. Practically, however, the online catalog field is not at the level of maturity that would make such standards sensible. We simply do not know what the most effective and useful displays look like and cannot agree on the most effective displays currently in use. That lack of consensus makes standardization nearly impossible. Standardization is unwise while there is clear room for major improvement through innovation.

THE COMMON COMMAND LANGUAGE

One almost-standard, which may finally reach adoption around the end of 1989, will begin to establish a common ground so that searchers can get started using new online catalogs. That standard, Z39.58, specifies a Common Command Language for Online Interactive Retrieval, usually abbreviated CCL.

CCL establishes a straightforward, nonlimiting syntax for command-driven interactive retrieval. It does not rule out menus and does not limit the scope of an online catalog. It does, however, establish a small set of command words to carry out certain functions common to almost all online retrieval systems. Three examples are FIND, DISplay and STOP.

CCL can improve access in two ways. First, users of one CCL-compliant system would be able to start using a different CCL-compliant system with little or no training and be reasonably sure they were doing the right thing. Second, CCL offers a good basic checklist for functions that should be supported in an online system — a checklist that includes functions currently not present in many catalogs.

CCL will be more visible than many other technical standards, but its primary goal is that of every technical standard: to establish a common ground, a clear shared understanding of some particular aspect of a material or system. When properly devised, the common ground of standards simplifies the work of everyone, but does not restrict ingenuity and improvement.

LEARNING MORE

NISO relies on the voluntary efforts of hundreds of librarians, publishers and information scientists, as well as the membership fees and supporting efforts of several dozen members. Librarians should be informed as to standardization activities; at some point, it might even be desirable to participate in standards review or development.* There are three specific steps to learn more about technical standards in librarianship and stay aware of current activity in the field:

1. You can subscribe to and read *Information Standards Quarterly*, ISSN 1041-0031, NISO's newsletter, $40 per year, for up-to-date information on what NISO is doing and on related standards activity.
2. Join the Library and Information Technology Association (LITA), a division of ALA that provides the *LITA Newsletter* and the informative column "Standard Fare," which has appeared in almost every issue for the last 10 years.
3. The best way to keep up to date about standards is to keep reading new material about them. One could also read *Technical Standards: An Introduction for Librarians*, published by Knowledge Industry in 1986. The book is available in paper from G.K. Hall; ISBN 0-86729-192-3, ISBN 0-86729-191-5.

* The NISO Office will welcome inquires and provide information. Contact the National Information Standards Organization, P.O. Box 1056, Bethesda, MD 20817.

HYPERMEDIA-ENHANCED SUBJECT ACCESS TO CHILDREN'S MATERIALS

By Florence E. DeHart, Professor,
School of Library and Information Management,
*Emporia State University, Emporia, KS**

"The revolution in information and instructional technologies," according to *Information Power, Guidelines for School Library Media Programs*, "provides unprecedented opportunities for improving access to information and ideas."[1] This article examines how traditional subject access to materials for children through the junior high level may be enhanced by integrating the browsing function provided by the hypermedia format with the retrieval function of online public access catalogs (OPACs).

Hypermedia programs make possible linkages that can serve as subject analytics.

HYPERMEDIA AND SUBJECT ACCESS

Traditional subject access to children's materials has emphasized access to items viewed as a whole, with less attention to the consistent provision of subject analytical entries. The availability of hypermedia programs, however, makes possible the creation of linkages that can serve as subject analytics for the targeted material and as a thesaurus embedded within the items.[2] This will become clearer as the following explanation of the hypermedia format unfolds.

Various definitions of hypermedia, a broader term than "hypertext," appear in the literature. Smith defines hypermedia in relation to hypertext, "A hypertext system allows users to *link* information together, thereby creating trails through associated materials....*Hypermedia* is similar to hypertext, but instead of linking just text, users can link to other media, such as graphic, video, spreadsheets, animations and voice."[3]

Information in hypermedia format is stored in nodes and related through the links, or pathways, between associated nodes. Nodes in the Apple *HyperCard* system, included without charge with the purchase of a Macintosh computer, somewhat resemble a file box of index cards. In *HyperCard*, a file is comprised of a collection of cards called a stack. The individual cards, or nodes, can hold text fields, graphics and digitized sound.

A "find" command enables users to search through fields in a stack for particular terms. If desired, the *HyperCard* "find" command may be replaced by other retrieval software. An example is HyperKRS, a retrieval software program used in conjunction with HyperIndexer, which indexes each word in a stack. The HyperKRS software from Knowledge Set Corporation makes it possible to apply such database techniques as Boolean operators to searching the index.

The links between associated nodes are designated on the cards by "buttons," or highlighted areas. These buttons mark a trail to another card or stack and can be activated by clicking a mouse, a procedure familiar to Macintosh users. Examples include *Hyper Christmas Card* and *Reading Fun*. *Hyper Christmas Card*, published by Bright Ideas, is based on the Apple *HyperCard* system, which consists of an Advent calendar with buttons accessible for each day of December until Christmas. *Reading Fun*, a *HyperCard*-based children's educational tool published by Preferred Publishers, Inc., incorporates buttons for several levels of difficulty so that children can learn at their own pace.

The hypermedia environment also permits the use of programs written in languages such as C or Pascal to create external commands and functions. For example, as noted by Cisler, developers have extended HyperTalk, a scripting language built into the Apple *HyperCard*, "to control CD-ROM drives and provide an easy-to-use interface for online information services such as DIALOG and Information Access."[4]

Fuller explanation of hypermedia systems may be found in the literature. One source is Gluck's *Hypercard, Hypertext, and Hypermedia for Libraries and Media Centers*, scheduled for publication by Libraries Unlimited in Fall 1989. This book provides rationale behind use of hypermedia systems along

* Appreciation is given to Darlene A. Schueler, Graduate Assistant, for her help with this paper.

with an overview of hypermedia products, and it places an emphasis on the Apple *HyperCard*. Another pertinent work is Ching-Cheh Chen's *HyperSource on Optical Technologies*, scheduled for publication by the Library and Information Technology Association, American Library Association, in fall 1989.

Items published serially include *HyperLink*, edited by Roger Wood and published by Publishers Guild Inc., a bimonthly magazine for "the Macintosh owner wanting to learn and do more with Apple's *HyperCard* software." Subscriptions are available for the magazine only or for the magazine and disk combination. *MacIntoshed Libraries 2.0*, edited by Bil Vaccaro and Edward Valauskas, is another source of practical information. It is available annually either in text format or as a *HyperCard* stack from Apple Library Users Group. A self-addressed label should be included with the request.

Children may add their own contributions.

HYPERMEDIA MATERIALS

Children's materials available in hypermedia format, referred to as "stackware," cover a wide range of subject areas. In some cases, children may add their own contributions. For example, one of the stacks in the *HyperCard*-based *HyperAtlas* contains a card for each country that is linked to the map of that country. Children can enter their own geographically based data and then access the data through the map interface. A source that may be consulted for additional stackware is *The 1989 Macintosh Product Registry*.[5]

Jostens Learning Corporation, formed by merger of Education Systems Corporation (ESC) and Prescription Learning Corporation, provides courseware in mathematics, reading, science, writing and language arts for grades K-12. ESC, along with *Encyclopaedia Britannica*, developed *Compton's MultiMedia Encyclopedia* in hypermedia format, including mouse-driven commands, provision for inquiry-based research and a bookmark feature allowing students to keep track of their paths through the encyclopedia.

Librarians, teachers and students may create their own hypermedia materials through the use of hypermedia software authoring tools. Activity along these lines is already evident. For example, the Binghamton (NY) city schools[6] and the Cincinnati (OH) Country Day School encourage student participation in stack development. In the Cincinnati school, a fourth-grade student created an Apple *HyperCard* stack on Hannibal "that incorporates maps, text, pop-up footnotes, the digitized sounds of elephants trumpeting, and charts illustrating the depth of rivers at the time of year Hannibal would have crossed them."[7] Libraries might form networks to share hypermedia materials. Elementary school librarians might adapt the network concept behind the UNITE project (Unified Network for Informatics in Teacher Education) in Kansas. Loretta Wood, Librarian, Olathe South High School, reported in a telephone interview September 5, 1989 that UNITE, sponsored by the University of Kansas under a grant from the Apple Corporation, will permit several schools to share materials in hypermedia format.

Persons creating materials in the hypermedia format should incorporate features that appeal to children, such as color, repetition and consistency. Ofiesh offers an exciting suggestion for creating adult material that could be adapted for use with children's materials. To paraphrase his example, picture an edition of a children's magazine on a CD-ROM disk "and that the computer reading the disk has the following colored squares: red for definition, blue for graphics and/or pictures, green for clarification, yellow for remedial learning sequence, orange for animation, black for voice synthesis and so on."[8] Imagine that the squares could be accessed by touching them, using the mouse or keyboard or even by calling out the color. Wouldn't it be nice if a child could touch the blue and orange icons on the screen with the cursor flashing under the phrase "dinosaurs" and get a dynamic graphic representation of dinosaurs roaming in their natural habitats!

In addition to Apple *HyperCard*, examples of authoring systems include *Guide* from OWL Hypertext Products, the first commercial hypertext authoring program for personal computers; HyperBase, a system based on the Prolog AI (Artificial Intelligence) programming language, which permits attachment of expert systems to guide students through appropriate pieces of information; and Brightbill-Roberts' HyperPAD, applicable for use in the character-based DOS environment. Standards are lacking for these and other hypertext authoring systems, with the exception of the Standard General Markup Language (SGML), which enables hypertext authors to create links that can be read by most text editors.[9]

Attention to copyright clearance is an essential component of the process of creating hypermedia materials. As pointed out in guidelines included in Apple's *HyperCard K-12* Special Supplement, creators should be aware that scanning and digitizing images or sounds and saving them to disk is a form of duplication that comes under copyright law. The doctrine

of "fair use" familiar to librarians leaves room for interpretation and requires that a number of factors be considered, including how much of the copyrighted work is being reproduced. Librarians planning to distribute, share, or sell a stack being developed from copyrighted materials should be certain to obtain written permission from the copyright owner and identify the sources in the stack they create.

The ethics of using external commands and functions, as defined above, demand particular attention, according to Stephen L. Michel.[10] The reason is that these are usually distributed as compiled code, and it is possible to have stacks containing externals without being aware of who created them. Michel suggests keeping a copy of the stack from which the external command was taken because it may contain information about the licensing and distribution of the command.

As hypermedia materials are created or acquired, they should be made accessible in the library's online catalog.

MAKING HYPERMEDIA MATERIALS ACCESSIBLE

As hypermedia materials are created or acquired, they should be made accessible to users by including them in the library's online catalog. A catalog that includes a library's whole collection and provides equal access to materials regardless of format, including computer disks, may stimulate children to use and benefit from the particular features offered by various formats. This philosophy of integrated access is supported in the recent American Library Association publication, *Cataloging Correctly for Kids*.[11] The *Guidelines on Subject Access to Microcomputer Software* could be adapted for use with hypermedia materials, according to which Library of Congress subject headings and Dewey Decimal classification numbers could be assigned.[12]

The use of subject headings will serve the traditional function of identifying the contextual framework of an item viewed as a whole, as illustrated by the following hypothetical example. Imagine that *Piggybook* by Anthony Browne had been published in hypermedia format.[13] This book treats role of family members without explicit mention in the text. No specific words in the text could be identified, therefore, to serve as a button to lead to related information. Thus, an appropriate subject heading treating the book as a whole would be necessary to make this concept accessible.

ONLINE CATALOGS IN HYPERMEDIA FORMAT

Children are becoming increasingly familiar with computer systems at home and in schools that may have second-generation OPACs. For example, the Wichita (KS) Public School System, is in the process of placing terminals in the roughly 100 schools in the System to access an online catalog through the NOTIS system, a mainframe-based integrated library system originating from Northwestern University.[14] However, according to Hildreth, "hypertext operations, which permit navigation throughout the database in many different directions and the dynamic definition of 'related areas and interests' have not been implemented in second-generation OPACs."[15]

A next step, then, would be to transform OPACs into a hypermedia environment so that browsing among text, graphics and sound could take place in the same manner as among Apple *HyperCard* stacks. This is exactly what the Whitman County Rural Library System in Colfax (WA) is doing, where Victoria Waters, a library volunteer, is undertaking the creation of a "HyperCard Kids Catalog."[16] The projected microcomputer-based catalog, funded through an Apple Library of Tomorrow Program grant, will incorporate annotations, reviews and children's comments about the material. The goal is to provide as much information about each item as possible. The catalog will be used initially in bookmobile service, including school visits, with plans for later use at the headquarters library.

This enhancement of catalog content may turn OPACs into learning centers.

Here is how the hypermedia catalog accommodates browsing. If a child has looked up "BASEBALL," and a retrieved annotated card makes reference to the indexed name of a famous baseball player, the child could click the mouse on that button (highlighted name of player) to obtain citations to biographical material, or, if desired, a biographical sketch, picture or voice excerpt. Other suggestions for linking materials may be found in *The Web*, a journal published by the College of Education at Ohio State University devoted to reviewing books around themes, or in *Creating Connections: Books, Kits and Games for*

Children: A Sourcebook, which suggests ways for interrelating the curriculum through networks of resources.[17]

The hypermedia environment also accommodates browsing in a classification mode. For example, the "zoom-in, zoom-out" function of The Intelligent Catalog (described in a video available from Bibliofile), which accepts local MARC files, could allow an apple to be "zoomed out" to an apple tree, or "zoomed in" to apple seeds. Each conceptual level could reveal the existence and location of related items in various formats. This enhancement of catalog content may turn OPACs into learning centers. Objectives of this new type of catalog will need to be thoroughly considered and placed into perspective with Cutter's traditional objectives for catalogs.

Hypermedia features can be added to OPACs already in place through the development of interfaces. School and public libraries using NOTIS, for example, might consider adapting one of the *HyperCard* interfaces originating from academic institutions, including MacPAC (Macintosh Public Access Catalog) developed by Mann Library at Cornell University,[18] and the MacNOTIS experimental system at Texas A&M University. Paul Leeper and Sherrie Schmidt expect to publish an article, "MacNOTIS: Taking Online Library Catalogs a Step Further," describing MacNOTIS. These may be used to replace the standard NOTIS command-oriented terminal interface with *HyperCard* features, including capabilities for downloading and printing.

THE CHILD AS FOCUS

Integration of the retrieval function of online catalogs with the browsing function of the hypermedia format cannot be expected to eliminate the frustrations children may face when searching by subject. Created hypermedia links may not necessarily correspond with thought patterns behind children's inquiries. Assistance and encouragement must take the form of "process awareness," viewed by Conklin as essential to the creation and use of hypermedia, for "the reader can and must constantly decide which links to pursue."[19]

As noted by Vandergrift, "hypermedia offers heretofore unheard of opportunities to gain insight into the ways young people perceive, process, and use information"[20] and allows us to trace how children follow trails through materials and make modifications. Results of these audit trails could provide feedback useful in creating semantic relationships among terms that, in the words of Molholt, "offer potential for viewing concepts in new and unanticipated ways, unlocked from rigid subject control."[21]

The hypermedia format could enhance the role of school library media specialists.

The required technology and its cost may appear to be deterrents to further development of hypermedia subject access. Some may even question whether children need this level of sophistication. However, Dede refers to the notion that "novices need less powerful devices than expert users" as a "myth of power." He believes that machines that are too limited will not be able to support the next generation of advanced learning environments.[22] An optimistic note is provided by Dede's prediction that "by the 1990s, schools

ADDRESSES OF SOFTWARE VENDORS AND USER GROUPS CITED IN THIS ARTICLE

Apple Computer, Inc., 20525 Mariani Avenue, Cupertino, CA 95014.

Knowledge Set Corporation, 888 Villa Street, Suite 500, Mountain View, CA 94041.

Bright Ideas, Inc., 87A Ocean Street, S. Portland, ME 04106.

Preferred Publishers, Inc., 5100 Poplar Avenue, Ste. 706, Memphis, TN 38137.

Publishers Guild, Inc., P.O. Box 7723, Eugene, OR 97401.

Apple Library Users Group, 10381 Bandley Drive, MS:8C, Cupertino, CA 95014.

Compton's MultiMedia Encyclopedia, brochure, Jostens Learning Corporation, 6170 Cornerstone Court East, San Diego, CA 92121-3710, 1989.

OWL Hypertext Products, 2800 156th Avenue SE, Bellevue, WA 98007.

HyperBase, Intelligent Hypertext, Cogent Software, Ltd., 21 William J Hts, Framingham, MA 01701.

Brightbill-Roberts, 120 East Washington Street, Suite 421, Syracuse, NY. 13202.

Bibliofile, The Library Corporation, One Research Park, Inwood, WV 25428.

will be able to afford the equivalent of today's advanced engineering workstations."[23]

The hypermedia format could enhance not only subject access to children's materials but also the role of school library media specialists as they exercise leadership in finding out how children learn, in creating responsive learning systems and serving as inservice specialists in teacher development programs. For the public library children's librarian, the sentiment offered by Cade applies, "You must maintain the public library system's role as a catalyst for change for each individual library and its users."[24]

REFERENCES

[1] *Information Power: Guidelines for School Library Media Programs*, prepared by the American Association of School Librarians and Association for Educational Communications and Technology, Chicago, American Library Association, 1988. p. 37.

[2] Schroeder, Terry, "Hypertext and the Embedded Thesaurus," unpublished manuscript, 1989.

[3] Smith, Karen E., "Hypertext — Linking to the Future," *Online*, v12, n2, March 1988. pp. 32-33.

[4] Cisler, Steve, "Micro Monitor: The Main Events of 1987 — How They Affect the Library World," *Online*, v12, n1, January 1988. p. 93.

[5] *The 1989 Macintosh Product Registry*, Vero Beach, FL, Redgate Communications Corporation, 1988.

[6] Telephone interview with Doug Green, Director of Computer Services, Binghamton (NY) City Schools, on September 15, 1989.

[7] "Welcome to HyperSchool," *Hypercard K-12*, Classroom Computer Learning, Special Supplement, Sponsored by Apple Computer, n.d. p. S-14.

[8] Ofiesh, Gabriel D., "The Seamless Carpet of Knowledge and Learning" in *CD/ROM: The New Papyrus*, edited by Steve Lambert and Suzanne Ropiequet, Redmond, WA, Microsoft Press, 1986. p. 308.

[9] Fiderio, Janet, "A Grand Vision," *Byte*, v13, n10, October 1988. p. 237.

[10] Michel, Stephen L., *HyperCard: The Complete Reference*, Berkeley, CA, Osborne McGraw-Hill, 1989. pp. 636-7.

[11] *Cataloging Correctly for Kids: An Introduction to the Tools*, Chicago, American Library Association, 1989. p. 41.

[12] *Guidelines on Subject Access to Microcomputer Software*, Chicago, American Library Association, 1986.

[13] Browne, Anthony, *Piggybook*, New York, Knopf, 1986.

[14] Telephone interview with Melvern Schroeder, Library Media Services, Wichita (KS) Public School System, on September 5, 1989.

[15] Hildreth, Charles R., *Intelligent Interfaces and Retrieval Methods for Subject Searching in Bibliographic Retrieval Systems*, Washington, DC, Cataloging Distribution Service, Library of Congress, 1989.

[16] Telephone interview with Victoria Waters, Whitman County Rural Library System, Colfax, WA on September 5, 1989.

[17] Cleaver, Betty P., Chatton, Barbara, and Morrison, Shirley V., *Creating Connections: Books, Kits and Games for Children: A Sourcebook*, New York, Garland, 1986.

[18] Coons, Bill, and Eng, Wing, "MacPAC: A Hypercard Communications Program and Front-End Interface to the NOTIS Online Public Access Catalog," *T.H.E. Journal*, September 1989. In press.

[19] Conklin, Jeff, "Hypertext: An Introduction and Survey," *IEEE Computer*, September 1987. p. 36.

[20] Vandergrift, Kay E., "Hypermedia: Breaking the Tyranny of the Text," *School Library Journal*, November 1988. p. 35.

[21] Molholt, Pat, "Research Issues in Information Access," *School Library Media Quarterly*, v17, n3, Spring 1989. p. 132.

[22] Dede, Christopher J., "Empowering Environments, Hypermedia and Microworlds," *The Computing Teacher*, November 1987. p. 24.

[23] Dede, p. 20.

[24] Cade, Roberta G., "The Future of Public Library Systems: A View From the Tenth Floor," *The Bookmark*, v45, nII, Winter 1987. p. 117.

INTERACTIVE AUTHORITY CONTROL

By Edward O'Neill, Senior Research Scientist, OCLC, Dublin, OH

The Library of Congress's distribution of machine-readable authority records, libraries' increased use of machine-readable bibliographic records, the growth of automated local systems and the recognition of the importance of database quality are generating renewed interest in authority control. Advanced authority control is one of the more promising means of enhancing the quality and performance of online bibliographic systems.

However, only certain bibliographic fields benefit from authority control, which must be combined with other methodologies to achieve comprehensive quality control. O'Neill and Vizine-Goetz[1] review the application of authority control techniques in the broader context of database quality. Currently, authority control applications are limited to names and subjects. This paper focuses on the application of authority control to subject headings.

Errors in assigning subject headings pose a serious problem.

NEED FOR CONTROL

Errors in assigning subject headings pose a serious problem by their magnitude, impact and variety. For subject access, errors cause different types of problems depending on the retrieval methodology employed. When the typical Boolean approach is used, records with erroneous subject headings are often irretrievable, resulting in decreased recall and precision. With browsable indexes, erroneous entries litter the display, making it difficult to identify the desired heading.

The majority of errors commonly found in subject headings fall into one of three general categories: (1) typographical and/or spelling errors; (2) invalid form of entry; and (3) coding errors (incorrect MARC tag or subfield codes). There are, of course, some errors that defy categorization.

Typographical and spelling errors are common in subject headings. For example, O'Neill and Vizine-Goetz[2] reported that 64 variant spellings of the word criticism were found in the subdivision, "History and criticism." Such errors propagate within a hierarchical structure. For example, these 64 misspellings resulted in 801 erroneous subject headings in 1,156 different bibliographic records.

Common form errors include: (1) inconsistent inversion; (2) incorrect use of qualifiers; and (3) inconsistent use of abbreviations. Form errors are frequent in subject headings. Many subject headings are inverted, meaning that the noun is brought forward, but the rules concerning inversion are not always clear. For example, the headings "Water, Underground" (instead of "Underground water") and "Drinking water" (not "Water, Drinking") are both correct. The inconsistency in the rules governing forms of headings results in confusion and attendant errors. Failure to supply a required qualifier or superfluous use of qualifiers are also frequent form errors. An example of incorrect qualification is "Programming languages", which should be "Programming languages (Electronic computers)."

Incorrect use of abbreviations is also common. Changes in the rules for using abbreviations and failure to follow the established practice cause indexing and retrieval problems. For example, it was once acceptable to abbreviate the subheading "Politics and government" as "Pol. & govt."; the incorrect abbreviation "Pol. & gov." is still found in many headings. Now any abbreviation in the entry is erroneous.

Incorrect MARC tag or subfield code errors are the third class of errors in subject headings. For example, the heading "History and criticism" was incorrectly entered as a place subdivision by prefixing it with the subfield code "z" rather than the "x" required for general subdivisions. It has also been incorrectly entered as a place subfield and a period subfield. Similar problems result from the incorrect use of field tags. Incorrectly assigning of the "650" tag identifies "United States" as a topical subject heading. Omitting the subfield terminator is also common. This can result in the heading "France — Description and travel" being entered as a single

subfield rather than as a main geographic heading with a general subdivision.

AUTHORITY FILES

Authority files are the common element in most approaches to control the quality and consistency of subject headings. Burger classifies online authority control systems into three classes: (1) systems with authority files completely independent and separate from the bibliographic database; (2) systems with authority files closely related to the bibliographic database but not linked; and (3) systems with authority files linked to the bibliographic database.[3] An additional fourth class can be defined for systems where the authority file contains provisions for error correction. This fourth class of authorities can support both error detection and correction while class 2 and 3 authorities are limited to error detection.

The differences between a class 2 and a class 3 file are whether the linkage to the bibliographic file is explicit or implicit. A class 3 authority file is an index containing pointers to each bibliographic record to which the authority entry has been assigned. Both classes require that all entries in the bibliographic file match a corresponding entry in the authority file. A good authority control system using either a class 2 or class 3 file can significantly improve the quality of the controlled fields by detecting most, if not all, erroneous entries. However, such systems can be rigid, slowing the cataloging process. Typically, they are capable only of detecting that an entry is invalid, leaving it to the cataloger to either revise the bibliographic record or to add a corresponding entry to the authority file.

ONLINE AUTHORITY CONTROL

Advanced authority control systems will include error detection and correction algorithms to overcome the rigidity normally associated with authority control. The system could provide the cataloger with as much information as possible rather than responding with "Invalid Entry" or other similar terse responses. For example, when a bibliographic entry is made that does not match an authority record, the authority control system could produce a short list of similar valid headings.

The correction aspect of an authority control system requires (1) a means of detecting the errors in the bibliographic records and (2) a means of correcting or assisting in the correction of the errors that are identified. In most current systems, detection alone is con-

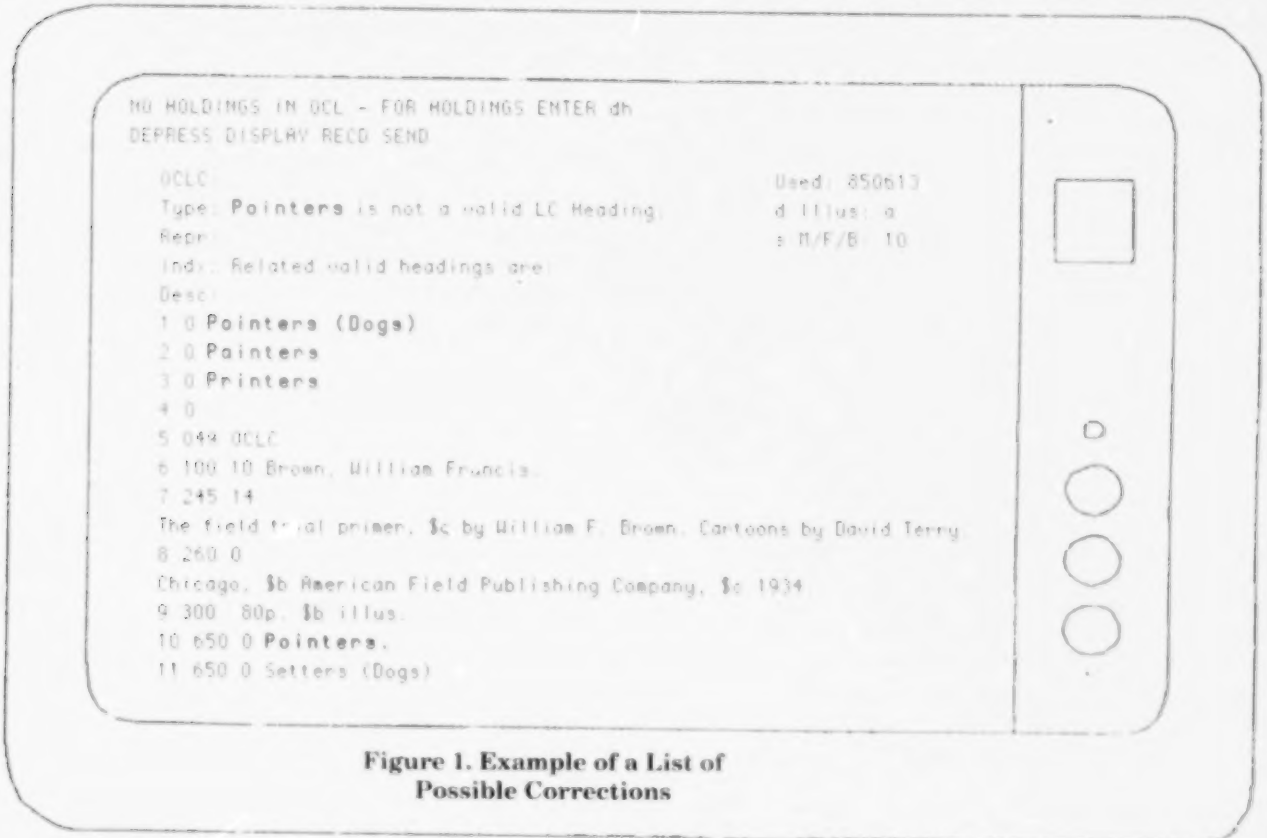

NO HOLDINGS IN OCL - FOR HOLDINGS ENTER dh
DEPRESS DISPLAY RECD SEND
OCLC:
Used: 850613
Type: Pointers is not a valid LC Heading.
d Illus: a
Repr:
s M/F/B: 10
Ind:: Related valid headings are:
Desc:
1 0 Pointers (Dogs)
2 0 Pointers
3 0 Printers
4 0
5 049 OCLC
6 100 10 Brown, William Francis.
7 245 14
The field trial primer, $c by William F. Brown. Cartoons by David Terry.
8 260 0
Chicago, $b American Field Publishing Company, $c 1934.
9 300 80p. $b illus.
10 650 0 Pointers.
11 650 0 Setters (Dogs)

Figure 1. Example of a List of Possible Corrections

sidered sufficient; the user is left to provide the correction. Error detection, however, is only half of the problem. To achieve high-quality, productive and satisfactory control, error detection should be followed by interactive error correction.

As noted by O'Neill and Vizine-Goetz, authority control is similar to the more general problem of spelling correction and is being influenced by the advances in that technology. Authority control can be viewed as a special purpose spelling checker that uses an authority file as the dictionary. If the standard definition of a "word" is changed to allow the inclusion of blanks and punctuation, an authority file can be viewed as a list of valid character strings. All common approaches to the detection and correction of spelling and typographical errors are applicable to authority files.

The advancements made in the spelling checker interfaces have improved the useability of interactive spelling checkers and most of that experience is applicable to authority control systems. For example, when the authority control system correctly identifies an erroneous entry, an error correction algorithm is invoked to correct the error. If the record shown in Figure 1 were entered, the error detection algorithm would identify the subject heading "Pointers" as invalid. The error correction algorithm should then generate an ordered list of possible corrections that might include "Pointers (Dogs)," "Painters" and "Printers" as possible corrections. These possible corrections could be displayed in a "pop up" box as shown.

The same authority control technology that is used to control the quality of cataloging can also be applied to validating queries.

The same authority control technology that is used to control the quality of cataloging can also be applied to validating queries. Most of the types of errors that are made during cataloging are also made by searchers. Queries with typographical errors rarely produce satisfactory results. This problem is particularly common for end users since they often lack keyboarding skills, are generally unfamiliar with subject heading structure and usually lack the experience to identify a bad search. Applying advanced authority control techniques that use error correction and cross-reference structures to queries would significantly improve both system performance and useability.

Large and diverse databases need authority control to satisfy the quality and performance requirements of new bibliographic systems.

CONCLUSION

Large and diverse databases need authority control to satisfy the quality and performance requirements of new bibliographic systems. Authority control systems are benefiting from major advances in related areas such as spelling checkers and interactive interface design. The next generation will be capable of providing interactive control over the development and maintenance of bibliographic databases in a productive, user-friendly environment.

REFERENCES

1 O'Neill, Edward T., and Vizine-Goetz, Diane, "Quality Control in Online Databases" in *Annual Review of Information Science and Technology*, Martha Williams, ed., v23, 1988. pp. 125-156.

2 O'Neill, Edward T., and Vizine-Goetz, Diane, "The Impact of Spelling Errors on Databases and Indexes" in *National Online Meeting: Proceedings*, Carol Nixon and Louree Padgett, comps., Medford, NJ, Learned Information, Inc., 1989. pp. 313-320.

3 Burger, Robert H., *Authority Work: The Creation, Use, Maintenance, and Evaluation of Authority Records and Files*, Littleton, CO, Libraries Unlimited, 1985.

DEWEY INTO THE 90s*

By Peter J. Paulson, Executive Director,
Forest Press OCLC, Albany, NY

The Dewey Decimal Classification is perhaps the oldest subject access system still in general library use. Its 20th edition having been published early in 1989, the 113-year-old brainchild of Melvil Dewey still enjoys a very robust existence.

Devised by Melvil Dewey in 1873 and first published in 1876, the DDC is currently used by an estimated 200,000 libraries in 135 countries and has been translated into some 30 languages. In the United States, DDC is used by over 95 percent of school and public libraries, and by 25 percent of academic and special libraries. Overseas, DDC is the classification of choice in 22 national bibliographies, including the national bibliographies of Australia, Canada, India, Italy, Indonesia, Jordan, Kenya, Pakistan, the United Kingdom and Zimbabwe. DDC numbers are included in machine-readable bibliographic records in the United States (MARC) and in the MARC tape services of 12 other countries.

From a slim volume of 44 pages in 1876, DDC has now grown to a four-volume work of 3,388 pages. This growth and the changes that take place in each new edition reflect the changes in the world around us — the growth in knowledge, the evolution of new knowledge domains or disciplines, changes in the kinds and types of materials libraries collect and changes in how people organize their understanding of the world.

Forest Press, now a division of OCLC (Online Computer Library Center, Inc.), is the publisher of the Dewey Decimal Classification, as it has been since Melvil Dewey first used the name as an imprint in 1911. Forest Press was "invented" by Melvil Dewey, because wily old Melvil recognized that his classification would be useful only if it was kept current and up-to-date.

Editorial work on the DDC is done at the Decimal Classification Division at the Library of Congress, under a contract with Forest Press. This relationship has been a particularly felicitous one for DDC users because it keeps the editors in close touch with the application of DDC numbers to current literature. The Decimal Classification Division of the Library of Congress is, in fact, the world's largest user of DDC, applying it to some 100,000 new publications every year.

The Library of Congress is the world's largest user of DDC, applying it to some 100,000 new publications every year.

Two events of great importance to Dewey users occurred in July 1988. Coincidentally, both events occurred on the same day.

On July 29 of that year, a machine readable tape containing the DDC 20 database was sent by the editors at the Library of Congress to INFORONICS, a Massachusetts-based library software and database company, to be used as the publishing source for edition 20 of the Dewey Decimal Classification. On the same day, Forest Press and DDC became part of OCLC.

In my opinion, both these events will be of great benefit to users of DDC. DDC brings to OCLC a widely accepted system for organizing knowledge, with strong international presence. OCLC brings to Forest Press resources and skills to transform the DDC database into computer-age products for DDC users.

The publication of DDC 20 in January 1989 marked a significant accomplishment in the history of the Dewey system, culminating 10 years of effort on the part of editorial staff, Forest Press and advisors. The new edition includes the previously published, revised Computer Science and Sociology schedules, one phoenix in the schedules (Music) and one in the tables (British Columbia). Many new features have been added to make the set easier to use: for the first time a manual is included (with references from the schedules); there are more summaries (which give an overview of a class or subclass); and more instructional notes and examples are included. The entire text

*Paper presented at an ALA workshop on DDC 20 in Dallas, Texas on June 23, 1989. Revised September 1989.

has been reviewed to clarify and modernize terminology, and the index has been revised and improved.

DDC 20 represents a new beginning, placing an increasing emphasis on user convenience. Even the size and typography have been redesigned to make the set more compact and easier to use.

DDC 20 was edited using a sophisticated microcomputer-based editorial support system and was published from a database carefully designed to produce both print and other products. The development of the DDC database provides the basis for some important steps in the future.

Four clear steps can be taken with this database. These major uses come immediately to mind:

- Generate additional print and electronic products
- Make DDC a part of a cataloger's integrated workstation
- Provide structure and vocabulary enhancement for subject searching in information retrieval systems, including online public access catalogs
- Organize knowledge in higher-level information retrieval systems, such as full-text, knowledge-based or artificial intelligence systems

The following sections describe each of these four applications in more detail.

PRINT AND ELECTRONIC PRODUCTS

From the point of view of publishers and users, the DDC database is similar to having the whole unabridged edition of the classification continually in "type-standing," up-to-date and virtually ready to be published at any time. This characteristic offers us the opportunity to publish new editions at intervals more convenient to users, to put updates in print more quickly and efficiently and to reduce the impact on libraries of what have been perceived to be the "massive" changes of new editions in the past. It also can cost-effectively generate subsets of the print edition when there is an identified need, such as the class 200 (Religion) reprint to be published by Forest Press in December 1989.

The DDC database, of course, can also generate electronic products. We are committed to publication of an electronic version of DDC 20 on CD-ROM disk, with advanced search and "windowing" features, and possibly with an expanded and enhanced index. Development of a CD-ROM product, including testing, is probably on a 12- to 18-month time line.

CATALOGER'S WORKSTATION

The idea of a cataloger's workstation has been under discussion in the profession for some time, and DDC would certainly be considered part of an integrated array of electronic tools in such a workstation. The development of an electronic workstation for catalogers and classifiers is an active research project at OCLC. The interest at Forest Press is in using electronic data to assist the cataloger in the tasks of bibliographic description and subject analysis, rather than merely to present print data in electronic form. The DDC database is already being used in that OCLC research. A proposed research project at another institution will examine the use of artificial intelligence (AI) programs to assist with the reclassification of collections when DDC schedules are revised. One writer has recently described an AI program that has successfully assigned DDC numbers to music scores.[1]

ONLINE PUBLIC ACCESS CATALOGS

A landmark study on the use of classification systems to enhance subject searching in an online public access catalog was completed by Karen Markey at OCLC a few years ago. The study compared an online catalog that had traditional subject heading, title and key word search capabilities with a catalog enhanced by browsing displays and subject keywords from the DDC. A major conclusion of the study was this: while the two methods of searching are relatively equal in precision, recall and search time, each method retrieved different relevant bibliographic items from the same database.[2] In recent papers, Karen Mandel, Pat Molholt and Maria Bates have each proposed radical new approaches to user access in online public catalogs, and Mandel and Bates have shown how some of these could involve classification.[3, 4, 5]

DDC is currently used by an estimated 200,000 libraries in 135 countries and has been translated into some 30 languages.

ADVANCED INFORMATION SYSTEMS

Melvil Dewey was always looking for new uses for the DDC; for example, he once authored an article on the use of DDC as an office filing system. The availability of the DDC database presents us with the opportunity to think of new applications for DDC in computer-based information systems.

The potential use of classification systems to struc-

ture and organize knowledge in full-text, artificial intelligence and other advanced information systems is mainly a matter of speculation at present. The availability of classification systems in database format opens the opportunity to enhance those systems with features that will be present in the database version only. Potentially then, a classification system may accommodate some of the requirements of advanced information systems without affecting the services and products provided to present print customers.

DDC has had an exciting past and faces an exciting future. We at Forest Press look forward to working with our users to develop the electronic products that they need.

REFERENCES

1 Schultz, Lois, "Designing an Expert System to Assign Dewey Decimal Classification Numbers to Scores," *National Online Meeting Proceedings — 1989*, Medford, NJ, Learned Information), 1989. pp. 393-7.

2 Markey, Karen, and Demeyer, Ahn, *Dewey Decimal Classification Online Project: Evaluation of a Library Schedule and Index Integrated into the Subject Searching Capabilities of an Online Catalog*, Final Report, Dublin, OH, 1986. 382 p. + appendices (Report Number: OCLC OPR/RR-86/1, February 28, 1986)

3 Mandel, Carol, *Classification Schedules as Subject Enhancement in Online Catalogs: A Review of a Conference Sponsored by Forest Press, the OCLC Online Computer Library Resources Center, and the Council on Library Resources*, Washington, DC, Council on Library Resources, October 1986. 24 p.

4 Molholt, Pat, "The Information Machine: A New Challenge for Librarians," *Library Journal*, v111, n16 October 1, 1986. pp. 47-52

5 Bates, Marcia, "Subject Access in Online Catalogs: A Design Model," *Journal of the American Society for Information Science*, v37, n6, November 1986. pp. 357-76.

CLASSIFICATION AND CATALOGING EDUCATION

By Judith Hopkins, Technical Services Research and Analysis Officer, University Libraries, The State University of New York at Buffalo

"The online environment changes the ways in which subject information is accessed and available."[1]

The changes in the ways in which subject information is accessed and available inevitably impact the educational process: not merely in what is taught or how it is taught but also who is taught, or to be more precise, who needs to learn.

According to Robert Hayes, cataloging is the fundamental basis of the theory of librarianship and information science. "It is here that information organization and structure become identified and the means for handling them specified. The structures embodied in cataloging are the basis for all subsequent retrieval, including reference. Even the substantive tasks in library management revolve around the structure provided by the catalog."[2] Ensuring effective and comprehensive bibliographic control is central to the functioning of any library or information center. Material that is not accessible is rendered useless; cataloging provides access. But "how do we assure that all librarians, whatever their specialty, have an appreciation of the importance of cataloging?"[3]

The focus of the literature cited in article is twofold: (1) the need for librarians not working in cataloging areas to understand cataloging principles and procedures and (2) the current state of education for cataloging and classification. With the increase in the number of libraries where some form of holistic librarianship is practiced (as it always has been in very small libraries) as well as the growth in use of machine-readable databases with their requirements for precision in both construction and use of entryways, the former need is acute.

CATALOGING EDUCATION AND THE NONCATALOGER

Holistic or general librarianship can be defined as a form of librarianship in which one person performs all the functions or is part of a team in which all functions (e.g., collection development, cataloging and reference) related to a particular subject area are performed. The importance of the cataloging function in this approach is still recognized, at least by some libraries. According to Gisela Webb, Assistant Director for Administrative Services at Texas Technical University, all librarians at Texas Tech do both public and technical services work. All new professionals are trained first in cataloging and then in public-oriented activities; they then do each 50-50. After a while the ratio can switch anywhere up to 80-20, with the mix depending on the library's needs and the individual's proclivities.[4]

With the proliferation of machine-readable databases, many noncatalogers are becoming aware of the need to understand how data are organized for effective retrieval. Any searcher of an online database that is not under rigid authority control quickly learns the value of authority control. Those involved in the implementation of an online public access catalog (OPAC) realize the need to understand the structure and content (and in some cases, the details) of the USMARC formats. Those involved in bibliographic instruction, either formally or informally, feel the need to understand why the OPAC works the way it does. Woodard and Golden find "there is increasing pressure on the reference librarians to improve their knowledge of cataloguing rules,"[5] while Laurie Smith indicates that a "solid foundation of cataloging rules, principles, and practices is needed before online access can be accomplished"[6] and recommends grounding in cataloging for automation and reference staff personnel.

A further illustration of the value of a background in cataloging principles appears in a recent article by Ferguson and Brundage in *College & Research Libraries News*. The authors, reference librarians at San Jose State University, describe the procedures the reference staff in their library had followed to provide a subject guide to the various periodical indexes the library owned. The authors then admit that the work would have been performed more quickly and more efficiently if they had called upon the expertise of the cataloging librarians in that library. To quote them, "Lack of cataloging and indexing expertise hampered the smooth flow of this project. Several problems with data standardization might have been avoided if we had consulted with members of our Cataloging or Serials Departments as we wrote the

data entry form.... Catalog librarians would have been able to warn us of...authority and database problems....The Cataloging Department could have been of great help in determining specificity levels for subject headings."[7]

Ironically, as the need for a more widespread knowledge of cataloging is becoming more evident, library schools have reduced the amount of cataloging required.

CATALOGING EDUCATION IN THE LIBRARY SCHOOLS

Ironically, as the need for a more widespread knowledge of cataloging is becoming more evident, library schools have reduced the amount of cataloging required. Saab and Lamson's surveys show that between 1966 and 1976 some 11.7 percent of schools dropped cataloging as a course required for graduation.[8] Kathryn Luther Henderson reports that of the 62 accredited schools providing Robert Hayes with data on required courses in the 1980s, 79 percent required a course in cataloging and classification.[9] In 1987 Walter M. High presented the results of a survey of the accredited library schools listed in the 39th edition of the *American Library Directory*. Of the 55 usable responses, 50 schools had a required course.[10] That is 91 percent of the responses, but only 79 percent of the 63 schools listed.

THE ROLE OF THE BEGINNING COURSE

What are the objectives of the beginning course in cataloging? Opinions differ. Nancy Williamson believes the objectives of this course are not to develop technical skills but to provide students with an understanding of the retrieval process, the nature of bibliographical and textual information, its documentation and representation, and the provision of access to it.[11] However, as Saye points out, those who work in small and/or specialized libraries will have cataloging as one of their many responsibilities.[12] These persons are not likely to have taken advanced courses in cataloging so, if they are not taught how to catalog in the basic course, they will not learn how to catalog anywhere in their library school program.

The objectives for a course determine its content. Those schools that believe the beginning course should inculcate technical skills will incorporate the practice of cataloging in the first cataloging course. Other schools will postpone that until the advanced courses taken by those intending to specialize in cataloging and will devote the beginning course to such matters as how to use the catalog and how to interpret it to others, how the concepts of bibliographic control underlie the structure of automated systems and discuss the technical services issues that have an impact on the library community in general.[13] This approach in its turn raises the question of how well students can understand these topics without some experience in the practice of cataloging.

In practice (presumably because of this difference in objectives) the contents of the required beginning course in the various library schools vary widely as Walter High's survey showed. Williamson generalizes that one-third to one-half of the typical beginning course will be devoted to aspects of subject control of information.[14] She defines subject control as "those systems and methods which control the access to bibliographic and/or textual data according to their subject content."[15] Mary Ellen Soper believes that "there is a trend in library schools to reduce the attention paid to descriptive cataloging, and increase education in subject analysis."[16]

Henriette D. Avram tell us that whatever else is taught, an understanding of the USMARC formats is basic.[17]

Some students will want to take more advanced cataloging courses. Of what should their cataloging education consist: How to make entries for a catalog? How to manage an existing catalog? Cataloging research? According to Francis Miksa the combination of a user-oriented information retrieval ideal with the new electronic technologies has prepared the way for a return to cataloging education in its broadest sense — learning the principles of creating bibliographic systems from scratch, principles that are not specific to any one system, agency or type of material.[18] This broad interpretation of cataloging education is demonstrated not merely by courses specifically on cataloging but also through a wide variety of courses relating to planning, designing and implementing of systems.

According to Williamson, the well-rounded course of study that professes to offer its students the opportunity to specialize in the area of subject control of information should have components that include the following topics:

- The retrieval process
- The process of analysis and indexing

- Indexing and retrieval languages:

  Controlled vocabulary versus natural language
  Linguistics and retrieval systems
  Authority control
- Classification systems; classification theory; classification in alphabetic systems
- Alphabetic systems including thesauri, their use and construction; string systems
- Abstracting and indexing methods
- Natural language processing, including automatic text analysis and artificial intelligence
- Standards, codes and conventions
- Multilingual systems.[19]

To that I would add a prerequisite of a good reading knowledge of foreign languages.

CONCLUSION

"Still I can not help thinking that the golden age of cataloging is over, and that the difficulties and discussions which have furnished an innocent pleasure to so many will interest them no more."[20]

We may be in the midst of another golden age with its inevitable impact on the teaching of cataloging and classification.

Charles Amni Cutter made his statement when LC began to issue catalog cards, but recent statements about the demise of cataloging with the growth of shared cataloging utilities resemble it. Just as the prolific literature of the last 80 years gives evidence that "difficulties and discussions" continue to give cataloging librarians "pleasure," so there is evidence that announcements of the death of cataloging are premature. The shortage of cataloging librarians is widely acknowledged[21] and has resulted in various studies, conferences and publications. We may, in fact, be in the midst of another golden age with its inevitable impact on the teaching of cataloging and classification.

REFERENCES

1 Wolner, Theresa J., "Subject Access Literature, 1988" in *Library Resources & Technical Services*, v33, July 1989. p. 240.

2 Hayes, Robert M., "The More Things Change..." in *Recruiting, Educating, and Training Cataloging Librarians: Solving the Problems*. [Papers presented at a symposium held March 10-11, 1989 at Simmons College, Boston, MA] Edited by Sheila S. Intner and Janet Swan Hill. Westport, CT, Greenwood Press, 1989. (*New Directions in Information Management*, n19) pp. 10-11.

3 *Ibid.*, Hayes. p. 8.

4 LAMA Systems and Services Section Management Practices Committee meeting, Monday, Jan. 9, 1989, ALA Midwinter Meeting, Washington, DC, January 1989. [Notes of Judith Hopkins]

5 Woodard, Beth S., and Gary A. Golden, "The Effect of the Online Catalogue on Reference: Uses, Services, and Personnel," *Information Technology and Libraries*, v4, December 1985. p. 343.

6 Smith, Laurie E., "Where are the Entry Level Catalogers?,"*Journal of Library Administration*, v6, n2, Summer 1985. pp. 483-4.

7 Ferguson, Margaret S., and Brundage, Christina A., "Innovations: Using Innovacq to Create a Subject Guide to Periodical Indexes," *College & Research Libraries News*, v50, June 1989. pp. 483-4.

8 Saab, Nehmat G., and Lamson, Merle E., "Cataloging: 1966 vs 1976," *Journal of Education for Librarianship*, v20, n2, Fall 1979. p. 111.

9 Henderson, Kathryn Luther, "Some Persistent Issues in the Education of Catalogers and Classifiers" in *Education and Training for Catalogers and Classifiers*, edited by Ruth C. Carter, New York, Haworth Press, 1987. pp. 5-6. (Also issued as *Cataloging & Classification Quarterly*, v7, n4, Summer 1987. p. 25ff.)

10 High, Walter M., News Editor. "Cataloging News" *Op. Cit.*, *Education...* pp. 167-195.

11 Williamson, Nancy J., "Education for Positions in the Subject Control of Information," *Op. Cit.*, *Education...* p. 61.

12 Saye, Jerry D., "The Cataloging Experience in Library and Information Science Education: An Educator's Perspective," *Op. Cit.*, *Education...* p. 30.

13 Frost, Carolyn O., "Preaching to the Unconverted: The Cataloging Educator's Challenge," *Op. Cit.*, *Education...* p. 196.

14 *Op. Cit.*, Williamson. p. 62.

15 *Op. Cit.*, Williamson. p. 60.

16 Soper, Mary Ellen, "Descriptive Cataloging Education in Library Schools, Using the University of Washington as a Specific Example" in *Education...* p. 49.

17 Avram, Henriette D., "Developing Catalogers for the Nation: Problems and Issues in Training Cataloging Librarians for the Future" in *Recruiting...* p. 302.

18 Miksa, Francis, "Cataloging Education in the Library and Information Science Curriculum" in *Recruiting...* p. 291.

19 *Op. Cit.*, Williamson. p. 63.

20 Cutter, Charles Ammi, *Rules for a Dictionary Catalog*. 4th ed., rewritten. Washington, Government Printing Office, 1904. (Quoted in *Charles Ammi Cutter: Library Systematizer*, edited by Francis L. Miksa, Littleton, CO, Libraries Unlimited, 1977. (The Heritage of Librarianship Series, n3 p. 201.)

21 "CCS Task Force on Education and Recruitment for Cataloging Report, June 1986" in *RTSD Newsletter*, v11, n7. 1986. pp. 71-8.

ASCLA PUBLISHES ANNOTATED GUIDE ON LIBRARY COOPERATION

The Association of Specialized and Cooperative Library Agencies (ASCLA) has published the first annotated compilation of key documents in multitype library cooperation. *Multitype Library Cooperation: An Annotated Guide to Working Documents* is designed to improve access to "working" documents — unpublished or locally issued documents, reports and studies created by multitype cooperatives, consortia, networks, systems and state agencies. The majority of these documents are unavailable through standard bibliographic sources.

Documents included in the guide cover topics such as the governance of multitype organizations, automation, interlibrary loan, document delivery, reference and information services, coordinated collection development, continuing education and other programs and services. The guide includes plans, studies, reports, requests for proposals, program descriptions, program evaluations, manuals, directories and training materials.

Material was gathered through a survey of 200 state agencies and multitype organizations listed in the *Report on Library Cooperation*. The 111-page guide includes 173 documents. All documents included in the guide may be obtained from the contact person or agencies listed.

Send orders to American Library Association, Order Services, 50 E. Huron St., Chicago, IL 60611.

CLASSIFICATION ONLINE AND THE QUEST FOR KNOWLEDGE

By Richard S. Halsey, Dean, School of Information Science and Policy, Nelson A. Rockefeller College of Public Affairs and Policy, The University at Albany

The attainment of knowledge is a personal matter, a psycho-social process not amenable to easy definition or physical measurement. It is "subjective" knowledge that is being referred to in most of what follows. The other kind, "objective" knowledge, is externalized in books and stored in libraries and databases. Success in its recovery is the usual measure of an information retrieval system's efficiency. Karl Popper, one of this century's most influential philosophers and historians of science, has provided a clear and imaginative definition of "objective" knowledge.

> When referring to human knowledge, I shall usually have this objective sense of the word 'knowledge' in mind. This allows us to think of knowledge produced by men as analogous to the honey produced by bees: the honey is made by bees, stored by bees, and consumed by bees; and the individual bee which consumes honey will not, in general, consume only the bit it has produced itself: honey is also consumed by the drones which have not produced any at all (not to mention that stored treasure of honey which the bees may lose to bears or beekeepers). It is also interesting to note that, in order to keep up its powers to produce more honey, each working bee has to consume honey, some of it usually produced by other bees.[1]

Pursuing this analogy a bit further, the bees are the creative artists and scientists, the beekeepers, academic institutions, foundations and industry; and libraries, apiaries hosting hives and honey for the mind. "Subjective" knowledge, or knowledge at the point of conception, is not data nor is it the comprehension or interpretation of data, nor is it judgement transmitted to others. Viewed this way, books and published nonprint products contain only raw material, being part of a static, classified world. Acknowledgment of the general public's consciousness of this stuff is reflected in the annoying "you know, you know" verbal twitch that pervades much private conversation and most celebrities' public pronouncements via the media. In contrast, creative inquirers are reconnoitering in a lightening limbo between ignorance and knowledge; if they already had the knowledge they sought there would be no reason to inquire. Yet they must be less than totally ignorant; otherwise there would be no recognition of need, no impulse to move towards knowledge. It is in this predawn domain between total ignorance and subjective knowing where the most potentially productive researchers strive to reach intuitively known unknowns. Their motivation as well as the needs of seekers for "objective" knowledge must be honored by the interface between user and database.

We need to distinguish between the processes of knowledge recovery and knowledge discovery.

Improvement of the recall of "objective" knowledge by implanting subject classification within the search mechanism has been the object of investigation by Marcia Bates, Karen Markey, Tefko Saracevic and many others. Bates has described an "online subject access system incorporating an end-user thesaurus as well as a variety of linkages among terms and document indexing of a sort that encourages awareness of other search terms, exploration of unanticipated possibilities, the making of mental associations and the development of a 'feel' for the system."[2] Within her hypothetical design, "to get into the system and to get going searching effectively, the searcher need only hit the side of a barn, i.e., any reasonable English language word or phrase should get the searcher started and linked to explanatory, guiding information to assist in the search."[3] Saracevic and Kantor, in an analytical summary of studies involving implantation of the Dewey Decimal Classification and Library of Congress Classification within online catalog retrieval systems, have noted that "the degree of agreement or overlap in human decisions related to representing, searching and retrieving of information

is relatively low — the agreement hardly reaches about one fourth or one third of the cases involved."[4]

Ironically, Saracevic and Kantor further discovered that "while the files which are searched are structured (indexed, organized) in a way that aims toward a clustering of answers, searchers seem to be working in a way that de-clusters them." Searchers tend to behave illogically, unbundling and shuffling prearranged data.[5]

Despite what we can learn about the logic or illogic of searchers, the development of systems to accommodate the search for already recorded knowledge by utilizing classification can be helpful as long as a loose and interactive structure exists that can overcome intrinsic bias in hierarchical systems and allow for an evolutionary approach and proper response to shifting states of knowledge, varying characteristics and changing orders of human priority. Such systems as PLEXUS employing artificial intelligence-based intermediary systems and the automated assignment of documents to faceted categories characteristic of BIOSIS, function well for searching in the sciences. Bibliographic classification-aided cueing with human intermediary support produce excellent payoffs in recall and precision. This is a substantial accomplishment because there is probably enough of this settled knowledge in an unsettled world to solve most social, scientific and technical problems if that information were brought together at the right time, or for the right people, or for the right problems. Nonetheless we need to distinguish between the processes of knowledge recovery and knowledge discovery. The latter impels personal action while the former tends to reinforce the status quo. Moreover, the established mono-hierarchial universal schemes and more organic and advanced poly-hierarchical and faceted systems encode and convey preset images of the world in their characters and paradigmatic structures. These images and hierarchies are culturally determined and not universally shared. Imposition of any universal system would be a form of mind meddling and manipulation on a grand scale.

Search effectiveness should not be evaluated solely in terms of recall and precision but in regard to its capacity to stimulate innovation through browsing without the formulation of specific requests. Needed are multiple dimensional and ranking modalities, nonlinear options to accommodate the resolution of anomalies. Generic, partitioned, oppositional, complementary and other functional relationships between subjects should be accommodated. The user should be able to employ faceted and synthetic classification schemes for special subjects such as rococo music, ribosomal RNA and radio astronomy and switch over to hierarchical pyramidal schemes for searching general subjects. Even with all of these sophisticated options and features, whatever systems emerge will inevitably exhibit a degree of suboptimization.

Electronic superhighways for knowledge threaten to override the library's counseling function.

Consequently, dependence upon expert human intermediaries will persist and probably even increase. Libraries must eagerly adopt this role or be lost on the side lines like the forlorn, orphaned and deserted road stands on Route 66. Electronic superhighways for knowledge are being built that threaten to override the library's counseling function. Fortunately for libraries and librarians, genuine researchers are not to be absolved from the obligation to study original text, and this will constitute a guaranteed stay of execution for libraries. As Don Swanson has observed, "citation trails, bibliographic coupling, cocitation techniques, and statistical measures of indexing similarity are antithetical to the idea of undiscovered connections...Searching the scientific literature is inseparable from reading it."[6] Scientific advance will continue to be the harvest of informed critical judgment, inspired conjecture, happenstance and hard reading.

Discovery is a wheel of becoming, a picking of fruit, a grasping at life, a walk in the garden of humanity's intellect with its many vistas beckoning the persona. The librarian is less knowledge engineer than expert orchardist; planting, pruning, picking and packaging the fruits of knowledge and human consciousness, codifying and prompting elemental impulses, the condition for each impulse being the absence of knowledge — the blindness of ignorance that can only lead to despair and darkness. One must admit that whatever systems are built can be but teaspoons to stir an ocean of information. Librarians and information scientists can only describe and interpret the spoor and symbols of knowledge, thereby enabling the end user to gain the privilege of sovereignty over knowledge, rule and informed action. Such is the essence of librarians' and information managers' latter-day mission.

REFERENCES

1 Popper, Karl R., *Objective Knowledge*, London, Oxford University Press, 1972. p. 286.

2 Bates, Marcia J., "Subject Access in Online Catalogs: A Design Model," *Journal of the*

American Society of Information Science, v37, n6, 1986. p. 365.

3 Bates, *Op. Cit.*, p. 368.

4 Saracevic, Tefko and Kantor, Paul, "A Study of Information Seeking and Retrieving. III. Searchers, Searches, and Overlap," *Journal of the American Society of Information Science*, v39, n3, 1988. p. 207.

5 Saracevic and Kantor, *Op. Cit.*, p. 208.

6 Swanson, Don R., "Two Medical Literatures That are Logically but Not Bibliographically Connected," *Journal of the American Society of Information Science*, v38, n4, 1987. p. 233.

NYPL RECOMMENDS NOTABLE CHILDREN'S BOOKS ABOUT THE BLACK EXPERIENCE

The New York Public Library has published the 10th edition of *The Black Experience in Children's Literature*, a seminal bibliography that recommends more than 450 outstanding fiction and nonfiction titles for children from preschool to junior high school age. Originally issued by the Library in 1946 under the title *Books About Negro Life for Children*, the guide was first compiled by renowned NYPL Children's Librarian Augusta Baker to catalogue a special collection of children's books she established at a Harlem branch library. Long since renamed, the book list has been revised and expanded to encompass the growing number of materials that center on the black child or interpret black history and culture for young readers. While these books have a particular significance for black children, they can also open a door for all children into a world of many cultures and traditions.

Noting a recent increase in the number of new titles, Christine Behrmann, the Library's Children's Materials Specialist, said, "The decline of the ethnic presence in children's books — which was so notable in the 1984 edition of *The Black Experience in Children's Literature* — seems to have been reversed. While we are not seeing the flood of volumes that we did during the late 1960s and early 1970s, a small revival is evident, with 50 more titles on this list compared to 1984."

The rise in the quality of children's nonfiction in the field is also reflected in the book list. Biographies have moved beyond profiles of sports figures to chronicle blacks who have been prominent in the arts, literature and politics. The selection of books about Africa has grown significantly, and the bibliography now offers such eclectic choices as the rich photographic portrait of *African Journey; Cooking the African Way*; and four biographies of Nelson and Winnie Mandela. The section on the Caribbean includes a greater diversity of titles, and books about African Americans span history from the American Revolution to Vietnam.

The Black Experience in Children's Literature is available by mail from the Office of Branch Libraries, The New York Public Library, 455 Fifth Avenue, New York, NY 10016 for $5. Shipping charge: 1-5 copies $1.00; 6-10 copies $1.25; and $1.50 for bulk orders.

APPROPRIATE USER INTERFACES FOR SUBJECT SEARCHING IN BIBLIOGRAPHIC RETRIEVAL SYSTEMS

By Charles R. Hildreth, Chief Consulting Scientist, READ Ltd., Springfield, IL

Online catalogs and CD-ROM retrieval systems with user friendly interfaces are now commonplace, but first things first. What is subject searching? For this discussion, let's limit our context to bibliographic databases and retrieval systems, including online catalogs. Apparently there is a lot of subject searching going on, even in online library catalogs. It seems to me, however, that we take the meaning of subject searching for granted. Perhaps this explains why there is so little discussion in the literature of the concept itself. When we do define it or talk about the phenomenon itself, we either 1) describe it in terms of what it is not (i.e., searching for specific known items), or 2) couch it in terms of its ultimate end product, for example, "searching for information on a topic," or "retrieving documents relevant to a query."

Clearly, the theological, *via negativa* characterization is not good enough for informed discussions on such an earthly matter. On the other hand, the output, product-oriented definition of subject searching, which usually carries with it the assumption that the answer to our question is obvious, is not good enough either. I will return to this point in a moment.

As a scientific discipline typically matures, early issues, problems and questions do not so much get resolved or answered as they get restated or refined. The pressing or interesting problems undergo reductive analysis or restatement from new perspectives. This process yields new and different questions, questions usually more specific than the earlier ones. We get more manageable research topics to chew on, so to speak, but we are also able to address aspects of a complex issue in greater depth.

In recent years some information retrieval (IR) system and online catalog (OPAC) researchers have been refining our initial question, "What is subject searching?," and have replaced it with a set of related questions:

- How do searchers actually go about seeking information on a topic in the larger context of the library, or elsewhere?
- What are the varieties of search behavior undertaken by subject searchers in our bibliographic retrieval systems?
- What are the various search aims or objectives of subject searchers?
- What sort of search and display interactions are satisfying to different subject searchers?
- What are the interactive search and display requirements for a variety of kinds of subject searching?

These research questions represent a shift in our thinking about information retrieval activities from a "product orientation" to a "process orientation." Previously our efforts have centered on ways of improving the output of online searches. Good results have been measured in terms of high recall (proportion of relevant documents that are retrieved) or high precision (proportion of retrieved documents that are relevant). Attention to the process of searching has largely been confined to training search intermediaries to formulate good queries and to search efficiently (i.e., keep costly connect time to a minimum) to produce the final best set of results to be handed over to the end user of the information. Many good suggestions and techniques have come from IR researchers for improving the retrieval (output) effectiveness of IR systems or the efficiency of database search specialists.

Unfortunately, most present-day online information retrieval systems, including most OPACs and CD-ROMs, reflect in their design and operation a partial, inadequate conceptual model of information retrieval activity. This model describes the exact or best-match, output-oriented approach of most IR systems. The model assumes the presence of a known, specifiable information need (or subject topic) as a start. Materials that are relevant to that need or topic are represented by index terms such as keywords or subject descriptors, and the need is represented in a well-specified query. These representations are then "best-matched" by the retrieval system (set in motion by the search specialist) to produce the best output set of retrieved materials or information.

This "known-subject need, best-match, end product-oriented" information retrieval paradigm accounts for only part of the subject searching story. It is conceptually inadequate for explaining a variety of informa-

tion seeking situations or for describing different actual subject searching behaviors. For these reasons, a number of researchers, among them Cochrane, Bates, Belkin, Hjerppe, Oddy, Marcus, Markey and Tague, have proposed other retrieval paradigms or conceptual models equally well-suited to guide the design of information retrieval systems. Bates has proposed the "exploratory paradigm" to describe unfocused information seeking and other forms of browsing.[1] The assumption is that browsing, a complex activity in itself, is a primary, frequent or preferred mode of subject searching for many individuals.[2]

If our emerging IR and OPAC systems are to more fully reflect and support actual searching behavior and a variety of subject searching styles and objectives, their design must be informed by models and user requirements analyses that closely describe these kinds of behavior and objectives. Whether the interaction and retrieval requirements of different types of subject searching are met by new systems will be a matter for ongoing investigation and evaluation.

Our user research tends to be "tree-focused" rather than "forest-oriented."

To begin with, it would be fruitful if we would consider that there are a variety of types of subject searching objectives and requirements searchers may bring to our retrieval systems. When better understood, these different aims and requirements may explain the different kinds of subject searching behavior displayed by individual searchers. The lesson is: watch the users in a number of information seeking situations, using a variety of information resources and retrieval systems. Our user research tends to be "tree-focused" rather than "forest-oriented." Armed with this knowledge (rather than assumptions), we will be better qualified to make proposals for specific search features and interfaces appropriate for different kinds of subject searching.

TWO APPROACHES TO SEARCHING IN PRESENT-DAY ONLINE CATALOGS

Two fundamentally different search approaches can be found in second-generation OPACs (See Figure 1). It is commonly recognized that key-word, Boolean searching (postcoordinating) is different than precoordinated phrase searching, but the difference I am addressing now cuts deeper across the OPAC interface territory. It is the difference between exact-match Querying and Browsing. These different search approaches are best understood by considering differences that may in fact exist in 1) users' searching objectives, 2) system query specification and query input requirements and 3) system output displayed to the searcher and subsequent interaction (if any) with the searcher.

Figure 1. Two online catalog search interaction approaches.

I. **QUERYING***
   - A. Phrase Matching (Text strings or controlled vocabulary)
   - B. Keyword Matching (Discrete words, with Boolean or proximity formulations)

II. **BROWSING****
   - A. Presequenced, linear, inflexible (Typically lists of index terms, headings, descriptors or brief titles)
   - B. Nonlinear, multidirectional, flexible (associative navigation, hypertext, exploratory "serendipitous" browsing)

* QUERY search requirements: Search aim/criteria known and can be expressed with relative precision and completeness.

** BROWSE search requirements: Search aim/criteria not specific, not known, and/or cannot be expressed in appropriate query/indexing language.

There are two kinds of query searching: phrase matching and keyword matching. A query consists of a term or terms (e.g., a character, number, word or words, or a phrase) and the specification, sometimes called the query "formulation," which defines how the component terms of the query are to be interpreted or related for matching purposes (e.g., word truncation, Boolean combinations, word adjacency). The matching function of an OPAC is the mechanism through which the retrieval software makes a comparison between index terms that represent documents and query terms that represent information needs or requests. The matching criteria are specified through the query by the user or applied automatically by the system. Query searching of either kind (often called just "searching," to distinguish it from an OPAC's Browse mode) utilizes an exact-matching function on the part of the system, regardless of the manner in which the matching criteria are specified.

In this all or nothing approach, documents (bibliographic records in OPACs) will be retrieved in

response to a search only if an exact match of the query is found. The query may consist of a precoordinated phrase (with or without truncation) or a postcoordinated Boolean expression of keywords. In either case the Query search matching requirements are precise and rigid. The process is purely mechanistic. The burden is on the searcher to enter terms that will match the entry (index) terms in the database and to specify appropriate proximity or term relationship logic. No match means no retrieval, as viewers of our silent OPAC screens witness too often. The search may fail (i.e., not identify relevant documents that are in the collection) unless the searcher knows or guesses the exact way the term (word or phrase) appears in the subject index.

In keyword, Boolean queries, the system's matching mechanism makes a binary (yes/no) split of the database into bibliographic records that conform exactly to the requirements of the query, and all the rest. Only the former are retrieved as "hits." Partial or "closest" matching operations are generally not supported in second-generation OPACs and conventional retrieval systems.

Query searching is an appropriate, useful search option when the aim of the search is specific, when the searcher knows precisely what he wants, and when this request can be expressed in the language of the database. Even in subject searching for books or articles on a topic the searcher may sometimes know his topic exactly and may be able to express it in the language of the system (e.g., the assigned subject headings or descriptors).

Browsing in online catalogs can take many forms.

BROWSING

Browsing in online catalogs can take many forms. Typically, the system displays sequenced lists of terms, descriptors or brief bibliographic records for scanning by the searcher. Lists of index terms are usually presented in alphabetical order. The arrangement of brief citation records may be according to date, and some systems support short record browsing in shelflist order. Usually the only "navigation" option for browsers is to go backward or forward through the list in a semi-constrained, linear manner. Cross-references, if included, represent a way of jumping out of the sequence and over to related areas of the database. Hypertext operations, which permit navigation throughout the database in many different directions and the dynamic definition of "related areas and interests," have not been implemented in second-generation OPACs. Conventional browsing facilities assume the searcher has a vocabulary-selection objective to accomplish near the beginning of a search. They assist in identifying the correct form of a term and any related terms. Other forms of browsing, rare in today's OPACs, support related record/document discovery through nonlinear explorations of the database.

Browse searching is the most useful and preferred approach when the search aim is not specific (regarding, for example, discipline or topic, type of publication, level of treatment, perspective, etc.), the desired results are not precisely known in advance, or the correct terms for representing the user's query (which may be vague) are not known at the outset. One or more of these circumstances may be present in most subject searching activities. Searchers may often be motivated to use an OPAC or IR system because they recognize a gap or anomaly in their state of knowledge. Even if they have some sense of what information they need to resolve the anomaly, they can not precisely describe in a search query what they do not yet precisely know or know at all. As Gerrie explains, "A person cannot specify precisely what is needed to solve his anomaly, let alone attempt to couch that need in a way that defines a potential information source. It is reasonable to assume that there may be a mismatch between an information need that cannot be specified and the fundamental requirement of conventional IR systems for an exact logical expression."[3]

PRESENT-DAY OPAC SUBJECT SEARCHING SUMMARIZED

Subject searching in today's largely MARC-based OPACs is supported by a variety of exact-match query methods and browsing facilities. Exact-match query searches are basically of two types: 1) phrase searching on precoordinated subject headings, and 2) keyword searches on cataloging or indexing data in the bibliographic record (e.g., titles, notes, subject headings and their subdivisions). Keyword subject searches may be formulated as Boolean expressions to indicate the desired relationship between search terms, and some OPACs also permit explicit proximity designations in the formulation of the query. The trend is clear: most OPACs support both types of exact-match searching. However, these are hybrid systems that do not integrate or link these separate query options in any useful way during a search. It is usually up to the searcher to choose, through commands or menu selection, one query type or the other.

The syntax and mechanics of entering subject searches have been simplified, and the task of precise

search statement, query formulation has largely been delegated to the system software. Choice of subject query type is selected from a menu, or invoked by a pressed function key or simple typed command. Phrase searches are automatically processed as straightforward character string matches or word adjacency, same order matches. In most cases the match must begin with the leftmost significant word or character of the indexed entry. Keyword searches with more than one word are automatically processed as Boolean AND queries in most OPACs. Keyword searches may be targeted by the user or the system to one or more field indexes (e.g., title, notes, series statement, subject headings). Some flexibility is permitted in most keyword search OPACs. When searchers choose not to identify in their queries fields for subject keyword searching (or are not permitted by the system to do so), the trend is toward automatically searching both subject and title. "ANDed" keywords do not have to appear together in either the title field or the subject field to cause a match. Matching records would include those having one word in the title and another word in the subject heading.

Successful subject phrase searching in most OPACs still requires an exact match on at least the initial, main portion of the subject heading in the catalog record. This usually requires a perfect match on a Library of Congress Subject Heading (LCSH). However, if a LCSH heading is entered as a search term and the heading has been modified or replaced in the local catalog record no match will occur. In some OPACs, when no match occurs on the user-entered term, the system displays headings in the alphabetical neighborhood of the term. This may or may not help the searcher find a heading that expresses his search interest.

In present-day OPACs, browsing is even less developed and is generally provided as an intermediate or secondary stage in the subject search process. Some keyword OPACs offer the searcher an unpublished (on the screen) option to browse (scan) alphabetical lists of index terms or headings. Phrase search OPACs typically display these lists after a search is constructed and entered, even when an exact match occurs. Research has not shown whether, in the latter case, this helps or confuses the searcher. A few OPACs are beginning to include cross-references and related terms in these browsing displays. This will make them more useful browsing tools for certain kinds of subject searching.

There has been a limited introduction of bookshelf-style browsing via the OPAC display screen. Unfortunately, the small amounts of book data typically displayed will not support the kind of actual shelf browsing practiced by library patrons in open stacks. Moving on at the terminal to additional online information for a selected document, then returning (if permitted) to the browse list, can be cumbersome. There is no good reason to restrict shelf list browse entries to one or two lines of display data, although this is usually done.

LIMITATIONS OF THE CONVENTIONAL MODEL

Conventional IR systems place the burden on the user to reformulate and reenter searches until satisfactory results are obtained. This is typically the case with second-generation online catalogs as well. This approach assumes, however, that the user knows what she or he wants and can describe it in the language of the database being searched.

Hjerppe quite correctly rephrases this problem as the fundamental paradox of information retrieval, "the need to describe that which you do not know in order to find it."[4] Even the best second-generation catalogs do little to help the user transform an information need to explicit descriptions of the information understandable by the system. Nor do these catalogs lead the user from "found" information to related, linked information that has not yet been discovered. It is unrealistic to expect our catalog users to know in advance the structure and language of our library databases. It is equally unrealistic to expect online catalog users to be proficient in the various search approaches and techniques *before* they engage an interactive system in the retrieval process. Hjerppe reminds us that humans are much more adept at recognizing something than generating a description of it.[5] Online catalogs could take advantage of this human facility by permitting requests such as, "Give me more like this!"

More informative, flexible, and intelligent browsing facilities can be provided in the online environment. It is time for librarians and OPAC designers to recognize that subject browsing may be a primary search activity in the user's quest to discover materials on a topic, or to discover unknown items of potential interest. This kind of search activity requires more than term selection browsing facilities, or even thesaurus-based automatic switching among related terms. Related document browsing and discovery can be facilitated in OPACs through richer precoordination in the database of multiple subject/topic clues found in bibliographic records (e.g., linking title terms with subject headings with call numbers, etc.) and by providing more search navigation options between retrieved and unretrieved (but linked) records, that is, record to record "jumping" at the discretion of the searcher (e.g., "Show me more books from this publisher." "What other titles are in this series?" "What documents cite this work?")

Today's OPACs also fail to provide more traditional subject search aids.

Today's OPACs also fail to provide more traditional subject search aids. For example, they do not support browsing in displays of classification outlines or schedules. Neither do they permit related term lookups in online subject thesauri or lists of subject headings. Classification schedule terminology and systematic subject information can lead the searcher to relevant records not likely to be retrieved by the traditional phrase and keyword subject search methods. It is seldom recognized that even Cutter favored this approach. Immediately following the famous passage where he lists the "objects" of the library catalog, Cutter proposes the means for achieving them. To enable the finding of books on a subject, the catalog needs the subject entry, cross-references, and a "classified subject table".[6]

Present-day online and CD-ROM catalogs come up short in the performance and usability areas because they do not:

- facilitate open-ended, exploratory searching by following preestablished trails and linkages between records in the database in order to retrieve materials related to those already found
- automatically assist the user with alternative formulations of the search statement or execute alternative search methods when the initial approach fails
- lead the searcher from successful free-text search terms (e.g., title words) to the corresponding subject headings or class numbers assigned to a broader range of related materials
- provide sufficient information in the retrieved bibliographic records (such as tables of contents, abstracts and book reviews) to enable the user to judge the usefulness of the documents
- rank the citations in large retrieval sets in decreasing order of probable relevance or "closeness" to the user's search criteria

BEYOND EASY BOOLEAN SYSTEMS

Improving present-day online catalogs and IR systems is a two-fold challenge: 1) making them more effective retrieval systems, and 2) ensuring that they are usable and satisfying to a heterogeneous population of end users — some trained, but many untrained; some knowledgeable in one or more disciplines, but many at the initial learning stages in a discipline.

Much advanced (post-Boolean) information retrieval research and theoretical analysis has been directed to improving the search performance and retrieval effectiveness of IR systems in controlled, experimental environments. While research has focused on system performance factors, actual human searchers have been excluded from most IR experiments. IR researchers and theorists generally have been critics of Boolean logic-based IR systems, and have experimented with a variety of alternative approaches that either attempt to ameliorate the shortcomings of classic Boolean methods (e.g., "extended Boolean" processing and "fuzzy-set" retrieval) or offer radically different, non-Boolean retrieval operations (e.g., query and document term vector processing and statistical, probabilistic retrieval methods).[7] A consensus seems to exist among IR theorists and investigators regarding the shortcomings of IR systems that rely solely on Boolean logic query expression and processing. Salton presents the following list of reasons why conventional Boolean retrieval methodology is not well adapted to the information retrieval task:

1. Formulation of good Boolean queries is an art rather than a science; most untrained users are unable to generate effective query statements without assistance from trained searchers.
2. Standard Boolean retrieval methodology does not provide any direct control over the size of the output; some query statements may provide no output at all, whereas other statements provide an unmanageably large number of retrieved items.
3. Boolean methodology does not provide a ranking of retrieved items in any order of presumed usefulness, thus all retrieved items are presumed to be equally good, or equally poor, for the user.
4. The Boolean system does not provide for the assignment of weights to the terms attached to documents or queries; thus each assigned term is assumed to be as important as each other assigned term. The only distinction actually made is between terms that are assigned (with an implied weight equal to 1) and terms that are not assigned (with an implied weight equal to 0).
5. Standard retrieval methodology may produce results that appear to be counter-intuitive:
   a. in response to an or-query (A or B or ... or Z) a record or document with only one query term is assumed to be as important as a document containing all query terms;

b. in response to an and-query (A and B and ... and Z) a document containing all but one of the query terms is considered as useless as a document with no query term at all.[8]

Online catalog research and design has been directed to making post-coordinate, Boolean, library retrieval systems easier to learn and easier to use than the commercial models used by trained intermediaries. Little attention has been given to the performance limitations of Boolean OPACs, and no university-developed or commercially available OPAC uses any of the advanced post-Boolean retrieval methods that have been tested with some success in the retrieval labs by the probabilistic and fuzzy-set retrieval theorists.

The shortcomings of second-generation OPACs and Boolean retrieval systems are now well known. There is no doubt that a vigorous dialogue between IR researchers and online catalog designers could lead to improvements in online catalogs and other IR systems intended primarily for use by the "everyman" end user rather than trained search specialists. Much is to be gained by a sharing of their separate insights and theoretical or design advances.

NEW MODELS OF INFORMATION SEEKING

I think we must recognize, however, that much of IR research and theory (perhaps its mainstream) is not relevant to the OPAC situation and is not particularly useful for advancing OPAC development into the third generation of systems. There are at least two reasons for this. First, the extended-Boolean, extended-vector, probabilistic and fuzzy-set retrieval models tested in artificial laboratory conditions (generally excluding the human variable) have not led to measurably large improvements in system performance (using recall and precision measures), and as Salton and others have recently concluded, these advanced retrieval models "may be too complex for practical use."[9] Secondly, the information retrieval paradigm assumed in most IR research may be the wrong model for describing the information seeking and search process situation and behavior of both scholars and general library and information system users.

The information system requirements of OPAC and IR system end users differ considerably from those of trained search intermediaries. The behavior of trained search specialists is product/output-oriented. Their aim usually is to produce a high quality list of citations or other references for the end user. The "quality" of the output product is measured by such limited variables as recall, precision, and search efficiency (minimizing the costs to the user of the retrieval process). The search topic or need is usually well-understood (a "known-subject") and well-expressed in advance of the search, and the search is typically processed in a highly structured, subject-specific domain.

AS WE OFTEN SEEK

The search situation is fundamentally different in the case of scholars and many others who use online catalogs for their own purposes. These searchers are an expanding heterogeneous community having a variety of information-seeking needs and behaviors. Furthermore, the document collections available to them for online searching are multidisciplinary in coverage and, at present, are poorly structured and indexed. Evidence indicates that the *process* of searching and discovering is more central to end-user searching objectives and satisfaction than the delivery of any predefined product. Most end users are not going after a specific "known item," nor do they have a well-defined output product in mind at the outset of their interaction with the online system. Researchers may wish to branch out into new disciplines or unfamiliar approaches to a problem. Typically, end users wish to discover materials on a topic of interest, and they seldom have or wish to present a precise expression of that interest. Both the expression of the topical interest and the interest itself may change dynamically during the search and browsing activity.[10]

At the beginning of the age of mechanized storage and retrieval, Vannevar Bush criticized the prevailing linear, rule-constrained retrieval ("selection") paradigm described by the conventional model of retrieval:

> When data of any sort are placed in storage, they are filed alphabetically or numerically, and information is found (when it is) by tracing it down from subclass to subclass. It can be in only one place, unless duplicates are used; one has to have rules as to which path will locate it, and the rules are cumbersome. Having found one item, moreover, one has to emerge from the system and reenter on a new path.
>
> The human mind does not work that way. It operates by association. With one item in its grasp, it snaps instantly to the next that is suggested by the association of thoughts, in accordance with some intricate web of trails carried by the cells of the brain.[11]

As a solution, Bush proposed his legendary Memex personal storage and retrieval system, perhaps the first hypertext system envisioned.

Unlike the linear, highly structured, logical search strategy approach pursued by the efficient search intermediary (human or machine), much end-user searching can best be described as exploratory, circuitous and, yes, *fully interactive*. The search process is likely to be largely a trial-and-error process, having no particular predetermined end or outcome. In trial-and-error exploratory searching, both the experience of the search process and its initial results can lead the user to new or altered information needs and may lead to additional materials or information of value.

Much end-user searching can best be described as exploratory, circuitous and, yes, fully interactive.

Miller and Tegler correctly point out that the "traditional methods of evaluating information searches — recall and precision — have completely overlooked this generative, creative aspect of a search. By evaluating the product and not the process, recall and precision limit our understanding of information searches and fail to measure them effectively."[12]

While it is clear that measures other than recall and precision are needed to evaluate the performance and effectiveness of interactive, end-user retrieval systems (e.g., measures of usability, browsing effectiveness, exploratory power, etc.), recall and precision have been the measures of "success" most often applied in evaluations of new indexing or retrieval approaches tested in IR experiments. The aim of much of this research has been to develop automatic retrieval methods that, given a particular document collection and a specific query (an expressed information need), would increase either the recall or precision value of the document set retrieved. This approach is of limited value in the OPAC situation where "illogically relevant" information encountered during the search process may be more valuable to the searcher than is logically relevant material.

CONCLUSION

Information retrieval research has pursued two directions in recent years in looking for solutions to problems presented by conventional Boolean IR systems. Two schools of thought can be delineated. The first school I prefer to call the "automatic" school. It has a strong grounding in probability theory. This school relies heavily on statistics-based computational or algorithmic approaches that, when they are implemented in retrieval system software, are largely invisible to the searcher.

The second school of investigation I call the "interactive" school. The approach of this group relies heavily on constructive, ongoing interaction with the system user to find means of improving both search results and the search process. Major theorists of this school include Oddy, Belkin, Brooks, Doszkocs and Marcus.[13]

The work of the "interactive" school of IR investigators holds more promise for the improvement of online and CD-ROM catalogs. This school's understanding of the nature of end-user searching and its associated problems in uncontrolled environments, and the need to constantly engage the user in the search process, more closely reflects the reality of the typical OPAC-user situation.

Some of the automatic retrieval aids tested by the first school (e.g., term weighting, vector processing, query expansion, document ranking) may prove to be useful in a supporting role, but I agree with Marcus' assumption that in most end-user searching it is "vital to bring human intelligence into play in an interactive, mixed initiative environment and that the human input is best induced when the human can easily perceive the nature of the (system's) matching algorithm."[14] Moreover, the user should not be viewed merely as the best judge of good query interpretations and processing by the system (i.e., the source of relevance feedback), but as one who enjoys engaging the system and who can be both creative and productive in using it, whatever the searching and browsing aims may be or become.

Online library catalogs and their users may indeed be viewed as "messy" or, at best, "fuzzy" indeterminate variables that present numerous problems for researchers and system designers. Yet, if progress is to be made, these variables must be understood as they exist in all their varieties. We need better descriptions of user search behavior in a variety of OPAC environments, and we need a more complete topology of search tasks and objectives from the perspective of the users. What do they do and why do they do it? What would they like to do if they had more flexible search and browse systems? With answers to these questions, we can more productively manipulate other variables (such as database structure, retrieval methods and dialogue styles and techniques) in experiments aimed at improving the overall search and discovery process.

REFERENCES

1 Bates, Marcia J., "An Exploratory Paradigm for Online Information Retrieval" in *Intelligent Information Systems for the Information Society*,

Proceedings of the Sixth International Research Forum in Information Science (IRFIS 6), Frascati, Italy, September 16-18, 1985, edited by B.C. Brooks, New York, North-Holland (Elsevier), 1986. pp. 91-99.

2 Hildreth, Charles R., "The Concept and Mechanics of Browsing in an Online Library Catalog" in *Proceedings of the Third National Online Meeting*, New York, March 30-April 1, 1982, Medford, NJ, Learned Information, Inc., 1982. pp. 181-196.

3 Gerrie, Brenda, *Online Information Systems: Use and Operating Characteristics, Limitations, and Design Alternatives*, Arlington, VA, Information Resources Press, 1983. p. 142.

4 Hjerppe, Roland, "Project HYPERCATalog: Visions and Preliminary Conceptions of an Extended and Enhanced Catalog" in *Intelligent Information Systems for the Information Society*. pp. 211-225. See also: Noerr, Peter L., and Bivins Noerr, Kathleen T., "Browse and Navigate: An Advance in Database Access Methods," *Information Processing & Management*, v21, n3 1985. pp. 205-213.

5 Hjerppe, *Op. Cit.*

6 Cutter, Charles Ammi, *Rules for a Dictionary Catalog*, 4th ed., Washington, DC, Government Printing Office, 1904. p. 12.

7 For a review of this post-Boolean AIR research see: Bookstein, Abraham, "Probability and Fuzzy-Set Applications to Information Retrieval" in *Annual Review of Information Science and Technology* (ARIST), v20, 1985. pp. 117-151.

8 Salton, Gerard, "The Use of Extended Boolean Logic in Information Retrieval" in *Proceedings of the Annual Meeting of the ACM SIGMOD*, Boston, MA, June 18-21, 1984, *SIGMOD Record*, v14, n2, 1984 pp. 277-285.

9 Salton, Gerard, "On the Use of Knowledge-Based Processing in Automatic Text Retrieval" in *Proceedings of the 49th Annual Meeting of the American Society for Information Science*, Chicago, IL, September 28-October 2, 1986, v23, 1986. pp. 277-287. See also: Bookstein, *Op. Cit.*, p. 141.

10 Lancaster, F.W., *Information Retrieval Systems: Characteristics, Testing, and Evaluation*, New York, John Wiley and Sons, 1968. pp. 181-182. See also: Oddy, R.N., "Information Retrieval Through Man-Machine Dialogue," *The Journal of Documentation*, v33, n1, March 1977. pp. 1-14.

11 Bush, Vannevar, "As We May Think," *Atlantic Monthly*, v176, n1, 1945. pp. 101-108.

12 Miller, Connie, and Tegler, Patricia, "Online Searching and the Research Process," *College & Research Libraries*, v47, n4, July 1986. pp. 370-373. See also: Swanson, Don R., "Information Retrieval as a Trial-and-Error Process," *Library Quarterly*, v47, n2, April 1977. pp. 128-148, and Harter, Stephen P., "Scientific Inquiry: A Model for Online Searching," *Journal of the American Society for Information Science*, v35, 1984. p. 114.

13 Marcus, Richard S., "Design Questions in the Development of Expert Systems for Retrieval Assistance"in *Proceedings of the 49th Annual Meeting of the American Society for Information Science*, *Op. Cit.* pp. 185-189.

Doszkocs, Tamas E., "CITE NLM: Natural Language Searching in an Online Catalog," *Information Technology and Libraries*, v2, n4, December 1983. pp. 364-380.

Belkin, N.J., Oddy, R.N., and Brooks, H.M., "ASK for Information Retrieval: Part I. Background and Theory," *The Journal of Documentation*, v38, n2, June 1982. pp. 61-71; and "ASK for Information Retrieval: Part II. Results of a Design Study," *The Journal of Documentation*, v38, n3, September 1982. pp. 145-164.

14 Marcus. *Op Cit.*

DISPLAYING MATERIALS IN AN INTERDISCIPLINARY ONLINE WORLD: NEW APPROACHES

By Joseph R. Matthews, Vice President, Marketing, GEAC Computers, Inc.

WYSIWYG! WIMP! Rather than taunts from one teen to another, these acronyms reflect the rapidly emerging computer technologies that are having, and will have, an increasingly greater influence on the display of information.

WYSIWYG (pronounced "whiz-e-wig") stands for "what you see is what you get." This acronym emerged from the word processing industry to distinguish those products that displayed text on the screen just like it would be printed on a page. WIMP (pronounced "wimp") is a computer industry acronym for "windows, icon (a picture to represent something), mouse and pointer." When most people think of a WIMP-based system, they think of the Macintosh, but WIMP-based software is also available on other brands of computers, including the IBM PC family of micros.

Users are beginning to demand that they be able to gain access from one workstation to ALL information on a subject, regardless of its format!

It is important, however, to note that WIMP-based systems actually require significant amounts of computer processing power to be able to display the information. Several years ago, when the "Scholar's Workstation" was just beginning to be talked about, developers of the concept of the Scholar's Workstation dreamed about the ideal computer as the "3M machine" — 1 MB of CPU memory, 1 Million pixels (or tiny dots) to be displayed on a screen (the greater the number of pixels, the greater the resolution of the image being displayed) and 1 MIPS (Millions of Instructions Per Second) of raw computer processing speed or power. Today, there are a number of machines on the market that exceed these specifications, though not all of them are in the price range the average computer user can afford.

Aside from the fact that the majority of computer manufacturers and software developers are moving quite quickly towards WIMP-based systems (as witnessed by the creation of the Open Systems Foundation), the reason WIMP-based systems are of increasing importance to libraries is that more information is becoming available in an ever-increasing variety of formats. And users are beginning to demand that they be able to gain access from one workstation to ALL information on a subject, regardless of its format!

The futuristic scenario as imagined by the technology gurus and others suggests that a college professor or student is able to gain access to the campus telecommunications network from home, office (and some even suggest, in the library) with the Scholar's Workstation. From the Workstation he or she is able to gain access to the library's collection of materials (emphasis on multiple types of materials, not just bibliographic data describing the monograph collection). The individual makes a few selections using the mouse and pointer and then enters the search. For example, the individual might be interested in the "role of music on the productivity of office workers." The library's online catalog might indicate that there were books on the subject, audio tapes containing various kinds of background and foreground music (*Musak*™), videotapes of documentaries and PBS television shows and recent articles from several journals in the library. The library's online catalog system contains links to various databases, some of which are contained on CD-ROM. One of the CD-ROM databases contains graphic images that are a part of some course materials for an upper division class, "Sounds in the Office Environment."

The WIMP environment would allow the user to make selections of materials that are of interest. The user could explore one or more of these information sources by having the system open one or more windows. From the window, the textual, sound, video or graphic material would be presented. Once completed, the user would be able to "close" the window and return to a previous window or keep the window open

and move on to other related information sources. Sound farfetched? Sound and video windows optionally available to complement bibliographic data! Is this feasible or simply Sci-Fi speculation? The technology does exist today to assemble various pieces of hardware and software to create a real true multimedia online catalog.

This same WIMP-based technology has obvious appeal for other areas within the library. For example, catalogers would love to be able to display an authority control file in one window while doing original cataloging in another window. Then, with a simple command, they would be able to move the desired entry from the authority control file into the original cataloging record without having to type the entry. Productivity increases, and the number of errors decreases. Similarly, an individual at the circulation desk could be charging books to a patron and move to a different window to pay a fine and yet a different window to place a hold.

The reasons WIMP-based technology is so attractive are that:

- Users are visually oriented
- Users are inherently multitasking (working on several things at once)
- The technology focuses the user's attention on the tasks to be done, not the technology itself
- Tools should model and mirror reality as closely as possible (thus, WYSIWYG).

How is information to be displayed so that it will be immediately recognizable?

All this sounds pretty exciting! Ignoring cost, and the library's multimedia catalog will have costs involved, there is a fundamental issue that must be recognized and addressed by the library community or else the future library multimedia catalog will only be a record setting disaster. The issue? How is information to be displayed so that it will be immediately recognizable? What icons should be used? Will the use of the system be so self-evident that librarians will not want to create "classes" on the use of the system? Alan Kay, a researcher at Xerox when they created the forerunner of the *Macintosh* and now Senior Fellow at Apple Computer, once suggested to a group of librarians and online catalog system designers that "a multi-media library catalog should be judged using the same set of three criteria used to judge a video game: 1) where to put the money, 2) how to play the game, and 3) how to win when playing the game."[1] If the designer of the game failed to clearly achieve all three design objectives, no one would play the game. Imagine if these three criteria were applied to today's online catalogs. Not too many would pass the test!

Most of today's online catalogs are either line- or screen-oriented bibliographic text-based systems. Yet, despite these similarities, great diversity exists when comparing online catalogs in terms of how they display information for the user. And, even though these online catalogs are only considered a unidimensional catalog, very few of these online catalogs do a good job of consistently displaying textual information to the user. It's not as if the library community needs to do basic research to better understand the issues surrounding the display of bibliographic information on the screen of a CRT terminal. Research abounds about good screen design and has even been summarized in a series of suggested guidelines.[2] Rather, the library profession must overcome other issues.

A good friend and colleague, Joan Frye Williams, has suggested that librarians are not buying good screen design.[3] Among the reasons suggested by Williams, six are key:

1. *Lack of user knowledge*. The buyer, the librarian (an intermediary), is buying a system for the end user, based on incomplete and inaccurate information about the users real needs.
2. *Management overload*. Coping with day-to-day events leaves little room for esoteric issues like good screen design.
3. *Verbal bias* of the library profession means that there is little likelihood that visual and graphic design issues like form will surface.
4. *An orientation to detail* means that the buyer's focus often is too narrow.
5. *NIH (Not Invented Here) or parochial vision*. Librarians are ignoring the research findings of other relevant disciplines.
6. *RFPs (Requests for Proposals)* have become a goal rather than a means to an end. Thus the RFP focuses quite narrowly on functional specifications with few words or concerns about good screen design.

Opportunities for new screen design and the delivery of information from multiple information sources will be offered with technologies that are just around the corner. Yet, the library profession must begin to come to grips with the importance of and need for good design. Otherwise, as Richard Saul Wurman suggests, the library profession may be a contributor to Information Anxiety.[4]

REFERENCES

1 Kay, Alan, "Bit Mapped Displays and Online Catalogs" in *Online Catalog Screen Displays: A Series of Discussions*, Joan Frye Williams (ed). A Report of a CLR Conference in Austin, TX, March 10-13, 1985. Washington, DC: Council on Library Resources, June 1986.

2 Matthews, Joseph R., "Suggested Guidelines for Screen Layouts and Design of Online Catalogs," *Library Trends*, v35, n4, Spring 1987. pp. 555-70. For suggested guidelines in the WIMP environment, see: Apple Computer. *Human Interface Guidelines*. New York: Addison-Wesley, 1987.

3 Williams, Joan Frye, *Automated Systems Screen Design: A Consumer Perspective*, presented at the LITA Pre-Conference on Screen Displays, New Orleans, LA, June 1988.

4 Wurman, Richard Saul, *Information Anxiety*, New York, Doubleday, 1989.

BIG BOOKS FOR NEW YORK'S PUBLIC LIBRARIES — Small New York Libraries to Benefit from New Grants

The New York State Library, in cooperation with the New York Library Association, is cosponsoring a special effort to help New York public libraries with limited resources acquire The Library of America series, the definitive collection of America's greatest writers. The effort is made possible with grants from the New York Council for the Humanities (New York) and the Andrew W. Mellon Foundation (New York). Packets of material describing the grant program were mailed to each of New York State's 739 public libraries on January 31, 1989.

Roberta Cade, Director of Library Development at the New York State Library, commented, "With this new grant, 96 of our libraries now have a rare chance to add substantially to their collections of American classics. We hope all libraries in small communities will move quickly to take advantage of this wonderful opportunity."

The New York Council for the Humanities, a nonprofit cultural organization sponsoring public programs in the humanities for adult audiences, has contributed $24,000 to this new, cooperative program. This grant will help New York libraries participate in a $2.3 million national grant program that awards funds for the acquisition of 60-volume sets of The Library of America. The new grant will assist 96 New York libraries by supplementing the matching grant program funded by the Mellon Foundation that has already aided over 1,500 libraries (including more than 91 in New York) acquire the award-winning series.

The program for New York's libraries is based on the original nationwide matching grant program whereby a library received $500 for the purchase of a 60-volume set of The Library of America when it raised a matching amount locally. The $1,000 purchased the series, normally costing $1,850. According to Dr. Jay Kaplan, Executive Director of the New York Council for the Humanities, "The funds from our council would be used to grant each qualifying library in New York an additional $250 toward the purchase of the series," so that only "the remaining $250 would need to come from local sources."

For New York libraries to be eligible for the $750 grant ($500 from the Andrew W. Mellon Foundation plus $250 from the New York Council for the Humanities), they must have an annual book budget that does not exceed $35,000 and be open to the public 20 hours or more per week. Branch libraries and libraries that are part of large systems are eligible.

For more information about the program or for an application kit, write to The Library of America, 14 East 60th Street, New York, NY 10022, or call (212) 308-3360.

INFORMATION — A CHANGING RESOURCE

By John J. Regazzi, President, Engineering Information, Inc.

In today's world information is an economic good that is literally being bought and sold. In fact, information products are now being packaged and dispensed in high-tech vending machines. A small company in San Jose, California, for example, recently announced the availability of their coin-operated microcomputer, which, for 15 to 25 cents, will access data from a compact disc built into the system and dispense information on anything from the financial markets to AIDS to drugs. These systems will soon be available in airports, libraries and drugstores.

Information products are now being packaged and dispensed in high-tech vending machines.

Of course, the notion of information as a commodity has been with us for some time. Recognition of the value of information as a resource has been growing exponentially in this century, particularly in the last 30 years. During this recent past, the worlds of publishing and communications have been transformed dramatically from a primarily print medium to a complement of electronic media. This transformation, however, represents much more than a change of form. In reality, the fundamental process of publishing is changing. The way information is used and valued by consumers, from researchers to business managers to high school students, seems to be evolving in myriad ways. How such changes in mass communication and publishing generally will affect the engineering, research and technical communities still remains to be seen.

This article will focus on the challenges confronting technical publishers, specifically those organizations providing research engineering materials and technical reference information. In order to focus particular attention on the engineering community, I will describe the specific changes occurring today at Engineering Information, Inc. (Ei), one of the pioneering publishers in this arena. Though specific to Ei, this case study is representative of similar information companies and some of the issues they are facing as well. As I hope to illustrate, Ei and other publishers are being influenced by the changes in their client communities, and, in turn, each is seeking to respond to these changes with new forms and methods of communication appropriate to this new environment.

Ei: SOME BACKGROUND

Over a century ago in 1884, John Butler Johnson founded the organization now known as Engineering Information, Inc. Frustrated by what he called "my ignorance of the engineering literature," Johnson proposed to the Association of Engineering Societies that it publish an "Index Notes" section in the monthly journal. He worked at his own expense and developed his own methods of indexing and reviewing up-to-date engineering information.

The company was quick to realize that computer databases being created to print the journal offered an entirely new means of disseminating technical information.

Johnson strongly believed that this systematic review of the engineering literature would not only help to improve him professionally, but also improve the communication of this information among other engineering educators and other engineers. These "Index Notes" eventually became The Engineering Index, which is now published monthly by Ei. The enterprise that began as a few records each month now includes over four million references and abstract summaries of the engineering literature pub-

lished in journal articles, conference papers and a variety of other technical publications. This collection has been described as the "foremost keeper of technical reference" in the world.

Ei's ELECTRONIC PUBLISHING

After almost 80 years of publishing *Engineering Index* as a printed service, Ei, along with several other organizations, developed the means to publish this technical information electronically. At first, in the mid-1960s, Ei's electronic publishing effort focused on the use of computer photocomposition as a more efficient means of typesetting the monthly issues and annual volumes of *Engineering Index*. In fact, Ei was one of the first publishing companies to introduce automated typesetting into production processes. The company was quick to realize that computer databases being created to print the journal more efficiently had even greater potential as a new publishing medium and offered an entirely new means of disseminating technical information on a timely basis.

In 1968, Ei offered COMPENDEX (an acronym for COMPuterized ENgineering InDEX) on monthly magnetic tapes, and a new electronic information service was launched. This service is heavily used today in universities, research institutions, corporations and government agencies that have their own information retrieval systems and need quick and sophisticated access to the technical engineering literature. These in-house retrieval systems also tend to be offered in institutions that provide access to a fairly large user community, such as an engineering school or industrial group, and in cases where the students, managers and engineers are encouraged to seek their own information.

In many ways, the COMPENDEX tape service of 1968 was a precursor to a much larger change in the publishing industry, i.e., the availability of the electronic online database. COMPENDEX became one of the first commercially available databases when it was offered online in 1973 by two database vendors, DIALOG and ORBIT information search services. Today, COMPENDEX is available through these services, as well as nine other similar services, making a vast amount of critically important engineering literature accessible to hundreds of thousands of organizations and millions of individuals worldwide.

Ei — A PROFILE TODAY

Engineering Information, Inc. is the only not-for-profit information service whose sole mission is to organize the international engineering literature through indexing and the creation of abstracts and to disseminate this information globally.

To accomplish this mission, Ei reviews nearly 5,000 engineering periodicals and conference proceedings annually, containing nearly 500,000 articles or conference papers each year. From this material, Ei selects, abstracts, indexes and converts to machine-readable form approximately 200,000 records each year.

Figure 1 illustrates the growth of EI's enterprise since its beginnings. The rapid expansion indicated for the past five years reflects, in no small measure, the unprecedented growth in the amount of engineering literature produced worldwide, as well as Ei's efforts to keep pace with these exponential increases.

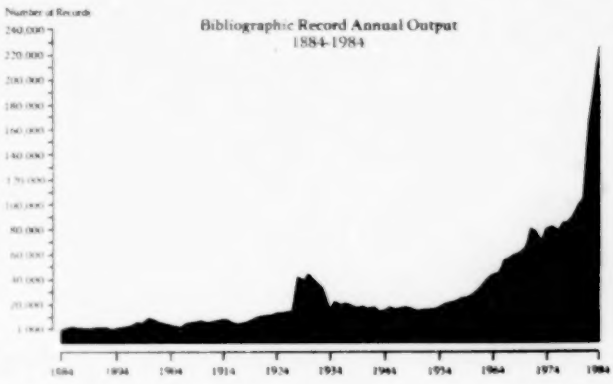


Figure 1

In order to disseminate this technical information to the widest possible constituencies, including developing nations, Ei has organized its efforts around three types of services: 1) printed publications, 2) electronic products and 3) information services. These products and services are designed to assist the practicing or research engineer in all phases of information access. For instance, Ei is involved not only in identifying a document that will meet a researcher's information needs, but also in supplying the source information as well. Last year, as part of its information services program, Ei instituted a document delivery service to provide researchers with the full text document of any published engineering article or paper. Orders can be accepted in writing, by phone or facsimile or electronically through major online ordering services. Other services, such as Ei's custom search program, are available to assist an engineer or organization in locating the most useful, pertinent information in the immense global network of technical references.

Notwithstanding the sweeping changes already effected by Ei and other publishers, the pace of information dissemination must continue to evolve if pub-

lishers expect to anticipate and meet the varied, complex and ever-changing needs of today's information seekers.

PUBLISHERS — IMPACTING THE FUTURE THROUGH THE APPLICATION OF TECHNOLOGY

Publishers are facing significant change in the publishing, communications and information industries. Publishers of abstract and index information, for example, are experiencing three major developments impacting the ways in which they provide service. First, publishers of secondary reference services are seeing an exponential growth in the primary literature. This trend places significant economic stress on these organizations to keep up with the filed as it grows. Second, this growth in the primary literature has been accompanied by a boom in the availability of secondary information retrieval products and services. There have never been more information goods available on the marketplace; moreover, the rate of new information products being introduced has increased in each of the last several years. Competition among information producers has never been greater, nor variety of information products more widespread. Third, and perhaps most significantly, information consumers are demanding more from their products and services. These demands tend to be for more current, specific and inclusive information content, coupled with more sophisticated and more diverse methods of accessing this information.

This technology will fundamentally reshape delivery systems for information goods and services.

Most publishers who have begun to meet these new demands successfully have done so through the utilization of new and more sophisticated technology. This trend will not change. The future direction of these technological developments will affect publishers and consumers alike. The expanding role of technology in publishing will not only affect the types of information products developed and made available, but perhaps more importantly, this technology will fundamentally reshape delivery systems for information goods and services.

Why will technology effect this change? I suggest two reasons: 1) the innovations in information technologies of the past two decades have radically reduced the time and cost of processing information; and 2) in reducing the costs of processing, information technology will permit an overall shift toward consumer and specific markets, replacing today's highly centralized and hierarchical delivery of information goods and services.

As illustrated in Figure 2, the cost of processing information today, compared with as recently as 20 years ago, has been dramatically reduced. If only a few basic indicators are reviewed, e.g., computational technology, storage and telecommunications, these will suggest substantial increases in speed and reduction in costs of information processing. Since 1970, there has been a 96 percent reduction in costs associated with computer processing units or the ability to process information. The savings realized by storing data on magnetic disks or optical discs are even more dramatic, exceeding in both cases a 99 percent reduction in cost. The ability to transmit information over telecommunication lines, though not as dramatic as either of the first two indicators, is still impressive with an 80 percent cost reduction for standard dial lines and a savings of 99 percent for leased lines.

Figure 2. Price/Performance Comparisons: 1970-1988.

	1970	1988
Computational Technology (CPU instructions per second)	$3	$ .12
Storage Technology (per million characters)		
Magnetic Disks	$2,500	$15
CD-ROM (Master)		$4
Telecommunications Technology (per million characters transmitted)		
Dial Access	$1,000	$200
Leased Line		$10

What these reductions illustrate most clearly is that the technology that launched the online electronic publishing industry of 20 years ago is now available to a much wider group of organizations — publishers and consumers alike. Most electronic publishing programs today still use a centralized approach, but with the application of low cost technology, a much more market-oriented approach is now possible. This market orientation will revolutionize the way consumers can request, access and receive information.

Centralized efforts, for example, tend to provide information goods through a single channel, chain or mode of distribution. In many cases two commercial firms are involved (i.e., a producer and a vendor); the producer of the information licenses a vendor to disseminate the information, and the vendor takes primary, if not exclusive, responsibility for those infor-

mation goods. Managerial decisions, not necessarily market forces, determine the design, content, price and availability of the product. Buyers do not select a supplier from a pool of vendors, but rather generally from a highly limited choice of select vendors.

Market forces will regulate that flow of information goods and services through the traditional economic gauge of supply and demand.

Markets, on the other hand, can direct the flow of information goods and services through the coordination of various individuals and organizations. Market forces will regulate that flow of information goods and services through the traditional economic gauge of supply and demand. In a market environment, information products may be easily linked to other products to meet complex market demands more efficiently. Thus, consumers may acquire specific information products from a variety of sources and make choices based upon the best combination of product attributes that meets their individual needs.

In this information market environment, publishers and consumers will be much more closely linked, and consumers will have a much wider range of choices regarding information products. Thus, some consumers may wish to pay 25 cents to find out the closing price of a stock, while others may have their own systems that will access a database they have acquired in electronic form from perhaps several sources and with their own software do trend analysis of this data. Moreover, if the consumer has questions about the data, he or she can contact the publisher directly through the same electronic system and additional information will be provided directly from the source.

CONCLUSION

In summary, there is a burgeoning demand for change in the publishing industry. Ei, as one example, has dramatically evolved since its inception but will need to continue to adapt if it is to meet the changing needs of engineers. As information technology develops and becomes more widely available, information goods will be delivered more directly to engineers. The form of the information will also change. Engineering information will not only be available in print and database form, but also in ways that engineers can put to use immediately on their workstations or personal microcomputers.

This market orientation, as it has been described, will benefit both publishers and consumers. Publishers should be able to supply information directly to their clients, thus being in a position to respond more quickly to changing needs. Consumers will have a much wider variety of options for information services, in many cases, being able to tailor a service to their individual needs, based upon the preferred aspects of several services.

Change is always difficult and, if not managed carefully, may provide a bumpy economic ride for publishers. By embracing technology, however, technical publishers will be able to offer better, more cost-effective and, most importantly, higher quality information goods and services to their clients.

FUTURE OF SUBJECT SEARCHING

By Richard T. Sweeney, Dean of Library Information Services, Polytechnic Institute of New York, Brooklyn

This article speculates about the future use of subject cataloging within sophisticated search engine technologies where the following man-machine interaction could take place.

- Do you have any books or articles on the use of CD-ROMs and LANs?
- There are two different books, each with one chapter, and 15 journal articles on Compact Disk-Read Only Memory and Local Area Networks. The text is being reviewed for more information. Three of the journal articles you have previously viewed, and two of those are saved in your file. One of the articles was written by John Doe, who has written several previous articles you have read. Do you wish to narrow this topic to more specific information on individual current products, theoretical articles, or reviews of the research and development? Do you wish to save these articles to your review file?
- Let me see an abstract of the article by Doe and a listing of the abstracts for the articles that specifically mention Novell token ring network software. Save all of the articles that are not already in my file.
- Displayed are abstracts of the six articles that specifically mention Novell networks. The first and second articles have covered the subject most thoroughly. You might also be interested in a related article that has a detailed review of a new CD-ROM jukebox research project.
- Yes, the jukebox article sounds good. How did you know I might be interested in CD-ROM jukeboxes?
- CD-ROM jukeboxes are designed to work on networks with remote access, and your past research indicates that you are interested in the problem of remote access to multiple CD-ROMs. Furthermore, most of the prior researchers who have searched this topic have expressed a similar interest in jukeboxes.
- Thanks. May I see a picture of the jukebox and its specifications.

THE GOAL

The goal of the search engine is to create a seamless, intuitive and transparent subject search that meets the searcher's needs in the quickest, most economical and user friendly manner. Such a system is seamless and transparent because the searcher is not aware of movement between databases, the command structures of different databases, the expert system voice recognition programs, or even the expert system natural language programs all operating simultaneously. Likewise, the searcher is unaware of the powerful search engine that is translating the spoken or written words into the information that satisfies the searcher's needs. The system is intuitive since it mimics the way two human beings, the searcher and librarian, would interact in order to conduct such a search. Built into the powerful search engine postulated by this interrogation, is a sophisticated but nearly transparent subject cataloging system, that will become the subject cataloging system of the future.

Will subject searching be needed in the future?

Will subject searching be needed in the future given the developments in various new information technologies? If so, what improvements or changes in subject searching can be expected? Which technologies will have an impact upon the future of subject searching? What does the librarian or library administrator need to know to plan effectively for future subject searching? What will be the demand for highly skilled subject catalogers?

THE DEFINITION

Subject cataloging has been defined as the use of a standard list of headings, composed of terms chosen to categorize the content or themes of individual cata-

loged documents (books, articles or other documents). Subject cataloging is more broadly defined in this article to include the cataloging of information units such as paragraphs, graphs, tables, charts, etc. within documents and the subject cataloging necessary to select and search across multiple databases. Information units have been labeled in some full-text databases as segments.

The key tasks of subject cataloging are maintenance of a subject heading list (authority file or thesaurus), selection of the appropriate subject headings for each document and selection of cross-references to other related subject headings. Maintenance of the subject authority file has the purpose of creating a controlled vocabulary that will link all possible documents on the same subject. The user must be able to easily translate his/her own words into the proper or accepted authority subject term that will in turn help the user locate all related documents. The selection of the appropriate authority term(s) is concerned with choosing terms that are unambiguous, most descriptive of the subject, stable and consistent and that best enable the searcher to find all the available records on a subject. Appropriate cross-reference terms are selected and added that can direct the searcher to other authority file terms where all or some of the records are related to the subject matter sought.

Subject searching was invented to assist users in finding information or documents related to their particular need through a structured list of the concepts about the theme or content. It has been proven many times that the subject approach to cataloging is the most popular with users precisely because it gets efficient, quality results without requiring the user to know exact information such as the author, title, or ISBN about a document. Subject cataloging is very effective when combined with other searching tools as part of a powerful search engine.

All subject cataloging requires a controlled vocabulary of terms from which a cataloger can select one or more terms to describe the theme or themes of a particular document or document segment. The control of the subject authority file is critical to the success of the searcher because the file assures that a cataloger can't arbitrarily or accidentally create a new term for a document where existing documents on that subject are already linked to an existing authority file subject term. Searchers who search by synonyms, idioms or related terms should be directed by cross-references and/or by expert system assistance to the subject authority term best related to the need. Consistency in the assignment of controlled subject searching terms directly affects the quality and performance of a subject catalog. The assignment of subjects terms is a human-assisted task, not one done by a machine alone.

THE PROBLEMS

There are problems with subject catalogs. Subject cataloging can be inconsistent; the authority file terms can be unrelated to the users vocabulary; adequate cross-references can be lacking; archaic subject terms are often not removed or replaced; and quality cataloging can require a costly outlay for labor. Subject cataloging can be greatly improved to correct these and other problems by the employment of some of the new information technologies.

No adequate replacement for subject searching has yet been invented.

No adequate replacement for subject searching has yet been invented. Subject searching is likely to be an integral part of powerful new search engines as far into the future as anyone can see. Subject cataloging is such an efficient and effective search mechanism that it most certainly will remain even though it will change. With subject searching, the labor of relating documents on similar subjects is largely accomplished by the cataloger rather than by the searcher. A cataloger assigns the subject terms only once when cataloging the item, while each of the many searchers will be relieved of the effort of pulling together all of the related document records. Searchers, using the subject catalog, can improve the speed of their searches and therefore they can improve their personal productivity and satisfaction.

Searching features such as key word searching, proximity searching, right- and left-hand truncation, Boolean, wild card, frequency count, browsing and other similar features are powerful components of almost any search engine but cannot replace subject searching. All of these features can be used either in databases with controlled vocabularies (e.g., subject catalogs) or with uncontrolled vocabularies (e.g., full text). Users who search for documents by keyword in a database with an uncontrolled vocabulary may be able to bring together some of the needed records, but they will also bring in many extraneous, unrelated records. Many important records will be missed. Such searching features are an excellent way to search a full-text database with a subject catalog. These search features can translate the user's vocabulary (the stated or typed request) into the most appropriate controlled subject term or terms required to do the most thorough and precise search.

For example, a searcher may conduct a keyword search, as noted in the opening dialogue, for "LAN" in an uncontrolled abstract database. This may not reveal all of the relevant books or articles listed under the cataloger's authority term "Local Area Network." However, a clever searcher will then look at the authority subject terms assigned to the relevant document records that were found with the term "LAN." One of the terms assigned to one or more of the related document records would be "Local Area Network." The searcher can then conduct a search for *all* the relevant records with the controlled subject term "Local Area Network."

A list of document records produced from searching for the term "network" will produce many irrelevant document records, such as books on the subject of "television network." Unrelated records can be filtered in advance with a Boolean limiter "BUT NOT," *if* the searcher expects them. Unfortunately, the searcher may not be able to anticipate the unrelated records. Unrelated records will cause the searcher to unnecessarily spend time paring the list. Alternatively, *all* of the records that are linked to properly selected authority control subject terms will contain, by definition, *only* related records. Of course, the quality of subject cataloging and the searcher's selection of the subject term are often such that the search results are not completely precise and accurate.

FULL TEXT

Subject cataloging has represented, until now, subject terms (links) selected by some human cataloger. Links between subject-related documents have only recently been automated. In addition, bibliographic document records, rather than the full text of all the documents, have been the target database of subject links. The manual selection of subject links has made the implementation of full-text subject cataloging both extraordinarily expensive and prohibitively slow.

A full-text database such as LEXIS, a database of court and legal decisions, is an example in which the searcher must use the words that a court *might* have used. Unless the court used a controlled vocabulary, the searcher is forced to do the work of trying to think up possible terms and combine them in the proper fashion. This is work that a subject cataloger might have done in advance.

EXPERT SYSTEMS

Subject cataloging of both bibliographic and full-text databases will begin to change with two new technologies, automated hypertext systems and expert systems. Expert systems are computer programs (and sometimes hardware and/or firmware) that have been developed in various fields to duplicate the knowledge and decision making of human experts. Unlike conventional algorithms, expert systems use heuristic strategies to arrive at solutions. Conventional algorithms are programmed to cover every possibility. If a conventional algorithm is programmed correctly, the program will have a predictable, correct and finite number of solutions. An expert system will apply its expert strategy to problems that have too many possible solutions to program in advance every correct solution. However, expert systems, unlike conventional algorithms, will occasionally have unpredictable and/or inaccurate outcomes.

For example, a conventional algorithm might be programmed for Tic-Tac-Toe to win or tie by programming *every* possible move. Programming every possible move in chess, however, is impossible because the number of moves is too great. An expert system might be (and has been) used in chess to employ various expert strategies to win depending upon the moves of the opponent.

Expert systems work well in an application where the problem is too complex to program every possible outcome. Expert systems do not work well in environments where the number of strategies (or rules) are too numerous, not well understood, or rely on common sense. They are "idiot savant" machines that operate like an expert in a particular, very narrow domain of knowledge but cannot function at all in areas requiring the use of common sense or broad knowledge.

Expert systems will change much of the look and feel of subject cataloging and searching, although the basic principles of subject catalogs will remain the same.

Expert systems will change much of the look and feel of subject cataloging and searching, although the basic principles of subject catalogs will remain the same. There will still be an authority list of subject terms, assignment of cross-references and preassignment of terms to documents or segments at the time they are entered into the system. However, there will be additional work occurring in the background, and

the searcher will not have the assistance of an "expert" in the selection of the proper subject term.

SUBJECT SEARCHING EXPERT SYSTEMS

Searchers must first select the appropriate databases (or files) that are the most likely match with the searcher's request. Yet each database may, and often does, have a different thesaurus of subject terms. Database Selection Assistance (DSA) is the expert system process of translating the searcher's stated request into the selection of the most appropriate database(s). Searchers are concerned with getting the references and documents that meet what they *meant* not what they said (or typed); therefore, the choice of the database should be done in the background following the system or searcher's previously defined strategy for selecting appropriate databases. Of course, such work must be monitored by the searcher and interrupted if it is not producing the desired results. The searcher's personal strategy could then be modified to obtain better results the next time.

There are many personal and system criteria that must be used in the database selection decision, such as cost of searching a particular database, physical access to that database, coverage of the database, currency of the database and so forth. The use of a common subject catalog to many databases has already been established as a critical part of the process of conducting an efficient search. In short, the database selection process may be done by an expert system that would evaluate users' needs, system requirements and the available databases through the use of a good subject catalog to all the databases. This subject catalog database could itself be totally maintained and operated by expert systems.

It is likely that, 20 years from now, an expert system will automatically link the user from uncontrolled user words to the controlled terms of the databases to be searched. I have labeled this expert system process Term Selection Assistance (TSA). The searcher's selection of the best subject terms remains the most difficult problem in using subject catalogs. Often the authorized subject terms are not as current as the searcher's vocabulary. If a TSA expert system is told that a few records found in the search of an uncontrolled vocabulary, such as the title of a document or the words of a court decision, are relevant to the searcher's need, the system, based upon past experience with other searchers, can assume which controlled terms should be used.

Translating the searcher's personal vocabulary into proper search terms can be done by a TSA system that has a good knowledge of the databases, the database searching commands and the vocabulary of the field. The TSA will take all of the uncontrolled words in a record that are good matches with the user's need and, using some advanced algorithms that mirror an expert searcher, match those user words first to authorized subject terms and then to the records of related documents.

Today's microcomputer workstations are getting more powerful with larger memories that make the expert system Personal Researcher Assistant (PRA) a reasonable possibility in the future. The PRA is defined here as the expert system software that operates on a searcher's workstation that will assist that particular person in conducting searches more efficiently, more effectively, with less effort and with better results. The PRA can maintain and employ the personal idiosyncratic vocabulary and searching strategies of the user. It should be expected that the work of PRA, TSA and DSA appears to the searcher as one seamless system.

A Cataloger Assistance expert system is likely to become an assistant to the cataloger.

A Cataloger Assistance (CA) expert system is likely to become an assistant to the cataloger with authority to decide when and if controlled subject terms or cross-references should be added or deleted or at least will have the authority to recommend terms to the cataloger. The expert system might have the capability of automatically adding new subject terms if they are popular, consistent and sufficiently well defined to the subject authority file. For example, if enough users search for "LAN" but then find relevant documents under "Local Area Network," the expert system may recommend the term "LAN" to the subject cataloger as a cross-reference, i.e., "See LAN."

Consistency in the use of subject headings for describing similar documents can be greatly aided by automated assistance particularly when cataloging is being done by more than one cataloger over a long span of time. Automation is needed to assist catalogers in remembering the proper term, building cross-references, correctly typing entries and linking new subject terms to old documents and terms. While authority file usage is getting a boost from automation, growth has not been as large as expected on a local library level because of the heavy staff cost involved in implementing and maintaining such systems.

Systems that begin to automatically recommend or even update authority control terms and cross-references based upon the interactive queries of searchers promise to greatly reduce the cost of maintaining authority control files. Such software is in the experimental stage. When this software is developed for industry it should easily and quickly be adapted for use in libraries.

At this point in time and as long as the assignment of subject terms and the control of the subject authority file are primarily staff intensive activities, the use of subject headings to identify segments of the full text of documents is and will be prohibitively expensive. This will begin to change with the development of expert systems and other more conventional algorithms. Expert systems will soon begin to replace some of the cataloger's judgement in the assignment of subject terms and in control of the subject authority file.

The development of the subject authority file is gaining importance today because of the demand for more efficient and effective searching techniques in larger databases. Clients have increased demand for searching techniques that reduce their searching time and provide more accurate results, i.e. just relevant information, from ever larger and more numerous databases. Databases containing the full text of documents are inevitably much larger than comparable bibliographic databases. Therefore, users can be expected to demand much greater subject access to the full-text databases of documents than they have to the bibliographic databases. Subject cataloging of segments in full-text documents will increase very slowly until the storage costs of such indexes have dropped and until better software can speed up both the productivity and quality of cataloging.

Full-text searching of documents has been supplemented with hypertext, which is a software utility enabling the editor to create predefined links between text segments, images and other parts of multiple documents. Such links allow searchers to browse from one idea in one document to a similar or related idea in another document. Hypertext is being greatly advanced by the development of expert systems that can assist the editor in the creation of links. Until the time consuming task of building links is fully automated, it is doubtful that libraries will be much affected by hypertext.

When hypertext has been enhanced with the automatic assignment of links by an expert system, the impact upon libraries will be awesome. Hypertext links to individual segments are similar to terms in a subject authority file. However, in hypertext the searcher skips the intervening steps of actually having to select from a controlled list of subject terms and goes immediately to the linked text segment. Each automatically assigned hypertext link must have some controlled meaning, like a subject term, in order to be useful.

Image databases are another example of very large databases that require very good subject catalogs.

Image databases are another example of very large databases that require very good subject catalogs. Since image databases do not have text, by definition, the searcher must find other ways to search. We can expect that some day, as better standards in imaging begin to develop, predefined icons or patterns could become their own form of controlled thesaurus "terms."

Imaging retrieval has already started to become the most important technology for libraries for the next generation. Library research, development, services and user expectations for image retrieval are growing geometrically and are likely to continue to do so. Libraries will be in a contest with industry to develop new concepts for labeling, recognizing, sorting and locating visual subjects.

In summary, subject searching will be of even greater value in the near and long term with the explosion in numbers of databases, very large scale databases, full-text databases, image databases and hypertext databases. The quality and efficiency of subject searching will, of necessity, be greatly improved to work effectively in this new and more complex environment. Searchers can expect that subject searching will be easier, more powerful and effective and more intuitive and transparent as it is eventually automated with new technology such as expert system cataloger assistance, expert system personal researcher assistance, expert system database selection, natural language processing and voice recognition and response systems. Subject searching will be more dynamic, more reflective of current usage, richer in cross-references, cheaper to produce and maintain and will use more consistent terminology. Subject searching will continue to become an integrated component of a growing array of searcher productivity features built into powerful new search engines. Subject searching will be heavily integrated into hypertext systems so that a searcher will be able to travel between the full text and images within and between documents using controlled subject terms

that will automatically be assigned by an expert system.

THE REALITY

Subject searching is a concept that has lived for centuries. It has survived, independent of specific hardware such as the card catalog and the computer. Subject searching has survived because it is a concept that most closely follows the way mankind thinks, communicates and remembers ideas.

Subject searching is a concept that has lived for centuries. It has survived, independent of specific hardware such as the card catalog and the computer.

NEW YORK STATE LIBRARY ANNOUNCES NEW RESEARCH RESIDENTS

Carole F. Huxley, Deputy Commissioner for Cultural Education in the State Education Department, has announced the winners of the New York State Library's annual Research Residency Awards. The awards were established in 1986 to support individual scholarship and to promote the use of the State Library's collections in scholarly research. The 1989 Research Residents are:

- Dr. John M. Heffron, instructor, State University of New York College at Fredonia, for the preparation of a college-level "reader" that will give students access to primary documents relating to the history and development of New York State;
- Dr. Ronald W. Howard, associate professor of history, Mississippi College, for a study of patterns of childrearing in colonial New York, 1624-1776; and
- Dr. William A. Starna, professor, Department of Anthropology, State University of New York College at Oneonta, for research on the history and development of Indian law in New York State.

The term of each State Library Research Resident is one year. During this period, each Resident has access to the collections and services of the State Library and the support of subject specialists on the Library's staff. The State Library has a collection of more than 5.5 million items and is the only state library member of the Association of Research Libraries.

EXPERT SYSTEM INTERFACE SOFTWARE IN MICROCOMPUTER WORKSTATIONS

By William H. Mischo, Engineering Librarian, Beckman Institute Library, University of Illinois at Urbana-Champaign, Urbana, IL

INTRODUCTION

The last few years have witnessed the introduction of several important information technologies, including powerful microcomputer workstations, optical disk and advanced magnetic disk storage media, computer graphics technologies and the availability of nationwide and campuswide computer networks. These technologies will greatly influence all aspects of library information services, including the provision of enhanced subject access to library users.

Some of the early applications of these technologies center around the development of campus information systems comprised of scholar workstations (located both in libraries and on individual's desks) linked via a campus network to local and remote computing resources.[1] These microcomputer workstations provide access to heterogeneous information resources and, to function effectively, must rely on customized interface software to guide users into and between the various resources available. It has become clear that the role of the user interface in information retrieval systems is of paramount importance, and the design of intelligent interfaces to facilitate retrieval has become an area of increased interest.[2]

The University of Illinois at Urbana-Champaign (UIUC) Library has been developing the components of a scholar's workstation. This work has explored the design of a microcomputer library workstation and focused on the development of user interfaces to discrete information resources. The microcomputer workstation employs customized interfaces for the statewide Online Catalog, the BRS Search software and locally developed databases stored as microcomputer hard disk files.

ROLE OF MICROCOMPUTERS

Microcomputer workstations play an important role in the UIUC Library's online systems. They are used both as public terminals and staff workstations in Library sites. These microcomputer workstations employ interface and gateway software to facilitate access to the campus and statewide online catalog; to provide a gateway function to external, complementary databases; and provide access to bibliographic, referral and directory information databases stored in microcomputer hard disk and optical disk files. The graph in Figure 1 illustrates the distributed access offered by microcomputer workstations. The effective

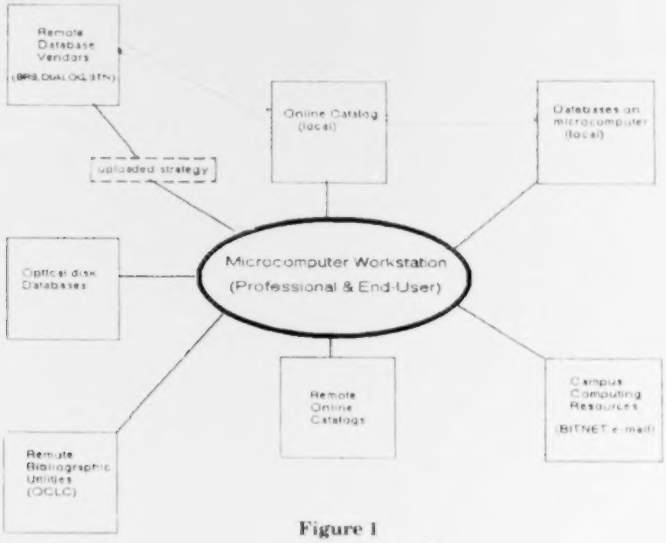


Figure 1

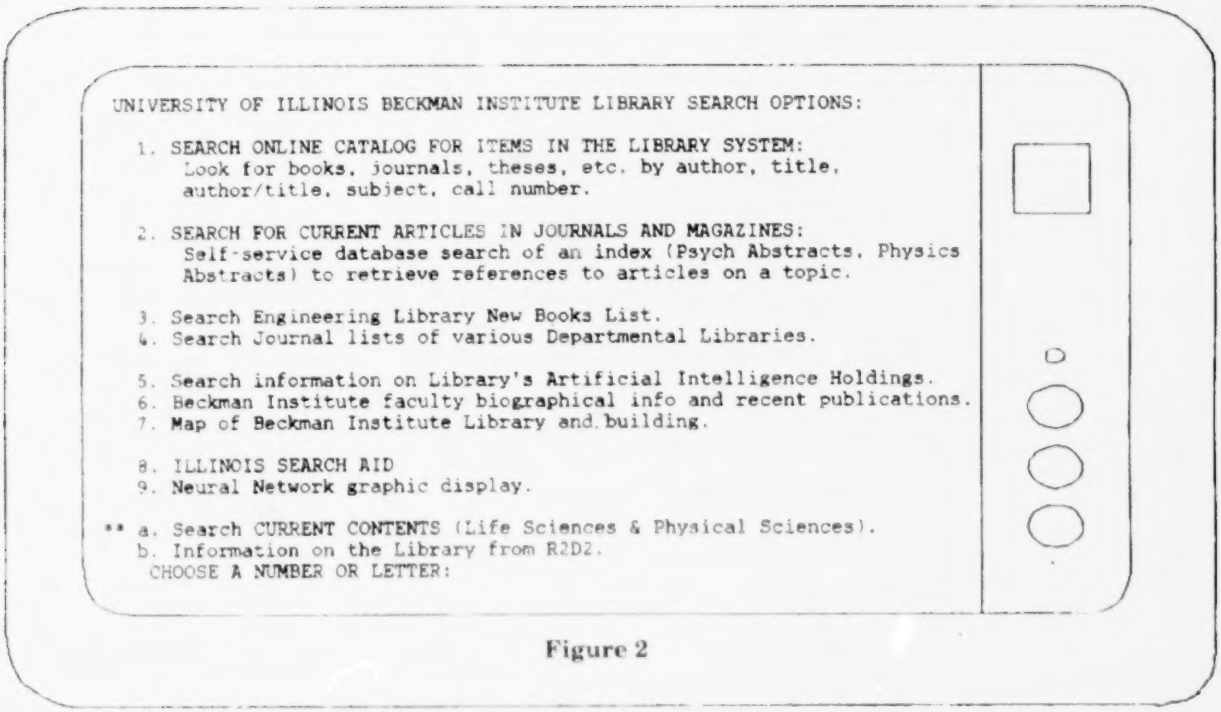

UNIVERSITY OF ILLINOIS BECKMAN INSTITUTE LIBRARY SEARCH OPTIONS:
1. SEARCH ONLINE CATALOG FOR ITEMS IN THE LIBRARY SYSTEM:
Look for books, journals, theses, etc. by author, title, author/title, subject, call number.
2. SEARCH FOR CURRENT ARTICLES IN JOURNALS AND MAGAZINES:
Self-service database search of an index (Psych Abstracts, Physics Abstracts) to retrieve references to articles on a topic.
3. Search Engineering Library New Books List.
4. Search Journal lists of various Departmental Libraries.
5. Search information on Library's Artificial Intelligence Holdings.
6. Beckman Institute faculty biographical info and recent publications.
7. Map of Beckman Institute Library and building.
8. ILLINOIS SEARCH AID
9. Neural Network graphic display.
** a. Search CURRENT CONTENTS (Life Sciences & Physical Sciences).
b. Information on the Library from R2D2.
CHOOSE A NUMBER OR LETTER:

Figure 2

utilization of microcomputers is made easier in the Illinois setting because the Online Catalog employs microcomputers as public terminals.

Both the Online Catalog interface and the BRS Search interface employ expert system techniques that mimic the actions of a reference librarian — particularly in the subject searching process — to navigate users through search sessions. For example, if a search results in too many matches, the user is prompted for additional terms to narrow the results. Searches that do not match a subject heading can be expanded to a title word search, and alternate subject headings identified in this way can then be extracted and searched. The BRS interface software uses a succession of BRS proximity operators to determine the best search results. Another important task performed by the interface program is that of distributing the processing by refining the search request at the microcomputer level, thereby limiting the number of transactions with the mainframe computer as the interface formulates the search arguments.

MICROCOMPUTERS USED AS PUBLIC WORKSTATIONS

The Library utilizes microcomputers at public service points to provide expanded access to information resources for library users. These workstations employ intelligent gateway software to provide access to multiple, discrete information sources, both remote and local. Figure 2 shows the opening menu of the user interface at the UIUC Beckman Institute for Advanced Science and Technology Library. (The Beckman Library serves as a showcase and laboratory for testing enhanced information technologies in the UIUC Library system.) This menu details the various components of the UIUC scholar workstation: new options are identified by letter to draw attention to their newness before being given a number. Available from this public terminal workstation are the campus and statewide Online Catalog (option #1), access to periodical index databases (option #2) and local databases and stored graphical displays on the microcomputer's hard disk (options #3-#b). Access to periodical index databases via option #2 is presently through dial-up to BRS, but the library will be purchasing the BRS Search software and mounting locally searchable periodical index databases.

A hypertext-like mechanism links data in local files with online catalog records.

The local data files stored on the microcomputer's hard disk are entered via word processing software or downloaded from the online catalog or remote

database vendors. These files are searched through locally developed database management software that features key word searching and browsing capabilities. These microcomputer-stored databases complement or expand on information available in the online catalog. A hypertext-like mechanism has been developed that links data in local files with online catalog records. Specifically, the keyboard arrow keys can be used to select call numbers from local files in order to display associated online catalog records in a window to extract up-to-the-minute holdings and availability information.

ENHANCED ACCESS TO PERIODICAL LITERATURE

Users of online catalogs exhibit an overwhelming enthusiasm for the technology and general satisfaction with search results. However, transaction log analyses and focused group interviews show that even satisfied catalog users are not necessarily performing effective and successful searches. Formal user studies of online catalogs have revealed that users experience serious difficulty with subject searching and that they expect online systems to provide enhanced access to a broader range of library materials, particularly periodicals. Providing enhanced access to individual articles in periodicals has been identified as an important focus of development in library automation efforts. Advances in online and optical technologies have spurred a dramatic increase in end-user searching activity over the last several years. Academic libraries now provide end-user access to periodical citation databases residing in: 1) integrated stand-alone systems such as augmented online catalogs; 2) optical disk retrieval systems; 3) discrete local online systems using search engines, such as BRS Search; and 4) remote bibliographic vendor systems, particularly the less expensive after-hours services. The UIUC Library is exploring these complementary techniques for providing enhanced access to the periodical literature.

USER NEEDS

Numerous end-user search services have been reported in the literature; a comprehensive review of end-user searching appeared in the 1987 *Annual Review of Information Science and Technology*.[3] Studies of the direct end-user searching of online catalogs, remote bibliographic databases and optical disk retrieval systems show that:

1. Users are enthusiastic about performing searches on easy to use, quickly learned, inexpensive search systems.
2. The searches performed are often not as effective as searches performed with intermediaries.
3. Search strategy formulation is the most troublesome problem for end users.
4. End users have considerable difficulty with Boolean logic.
5. End-user search services demand a significant investment of library staff time in training and assistance.
6. End users resist formal training sessions and the use of printed instructions, preferring CAI and direct one-to-one instructions from library staff or peers.
7. Many users search so infrequently that they remain naive users and require retraining.

In addition, much information about user searching behavior has been gleaned from online catalog use studies. These studies reveal that users experience the most difficulty with subject searching, and searchers rank subject search improvement as a high priority in suggested system changes.[4, 5, 6] In addition, users indicate a need for enhanced access to periodical literature. An in-depth statistical analysis of online catalog user questionnaire results suggests that a different type of user-system interface may be necessary for effective subject searching.[7]

These results have important ramifications for interface design and user training in online bibliographic systems.

EXPERT SYSTEM INTERFACE SOFTWARE

One of the most active areas of development in end-user information technology has been in search assistance software, or front-end interfaces, for aiding end-users in accessing bibliographic databases.

This one-stop-shopping approach provides library users with an "analytic" online catalog supporting access to multiple information resources from a single terminal.

Because of difficulties experienced by naive users in accessing the commercial database vendor systems

and the desire to link local and remote information resources within a single user interface, the UIUC Library undertook the development of customized microcomputer interface software for end-user searching of remote databases. This interface software utilizes some expert system search techniques to facilitate and simplify end-user searching of periodical index databases. This project has developed an enhanced microcomputer interface that permits searches of periodical databases to be performed as a search option within the local online catalog (see Figure 2). From an online catalog public terminal, users are provided an option to search for periodical information and then to link search results with local holdings and availability information. This one-stop-shopping approach provides library users with an "analytic" online catalog supporting access to multiple information resources from a single terminal.

The project has explored the use of expert system interface features in addressing the problems associated with end-user searching. The software facilitates effective end-user searching in online bibliographic systems. In particular, the interface assists users with search strategy construction and formulates the Boolean expressions in search statements for users.

The enhanced interface is being used at several departmental libraries. Over 1,600 end-user searches were performed in the last year. User questionnaires reveal that the system is regarded very favorably by users and is rated as easy to use. Transaction log analysis indicates that effective searches are being performed.

INTERFACE FEATURES

The software was developed under a campus-wide grant from IBM and a complementary grant from the Council on Library Resources. The remote searching feature is presented as an option in the online catalog menu. The project has also investigated linking retrieved periodical citations with local holdings and availability information in the online catalog.

The interface software features include:

1. conversational guided assistance and suggestive prompts in search strategy formulation
2. offline preparation and storage of search strategy
3. term selection from customized thesaurus term displays
4. software formulation of Boolean statements
5. automatic truncation of terms and stop word removal
6. replacement of user-entered terms with terms from a predefined substitution table
7. uploading of strategy and ranking of resultant search set combinations
8. software modification of search strategy, including limiting results to titles and controlled descriptors and using sophisticated search proximity operations, to optimize results
9. dynamic modification of search strategy based on user feedback

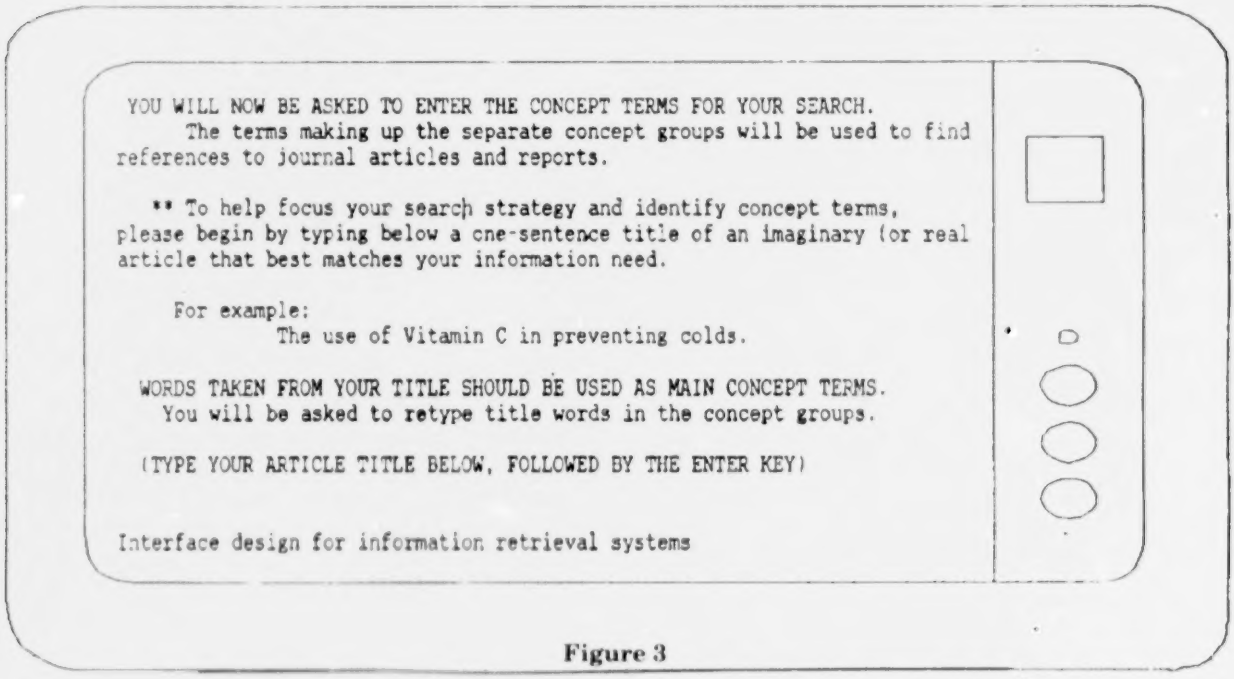

YOU WILL NOW BE ASKED TO ENTER THE CONCEPT TERMS FOR YOUR SEARCH.
The terms making up the separate concept groups will be used to find references to journal articles and reports.
** To help focus your search strategy and identify concept terms, please begin by typing below a one-sentence title of an imaginary (or real article that best matches your information need.
For example:
The use of Vitamin C in preventing colds.
WORDS TAKEN FROM YOUR TITLE SHOULD BE USED AS MAIN CONCEPT TERMS.
You will be asked to retype title words in the concept groups.
(TYPE YOUR ARTICLE TITLE BELOW, FOLLOWED BY THE ENTER KEY)
Interface design for information retrieval systems

Figure 3

61

10. selective downloading and printing of citations and search history

This expert system software approach incorporates search techniques employed by trained searchers and seeks to expedite the search process for end users. The user is guided through the search process by a series of menus, suggestive prompts and software manipulations that construct and modify the search strategy. Figures 3 and 4 show screens from an example search as it is being constructed. A more detailed sample search can be found elsewhere.[8] The remote access version of the interface software controls the search process, minimizing connect time and limiting the number of references printed to 25 per search session, although many users perform multiple searches. This results in essentially fixed cost or upper-limit costing on the search service.

EVALUATION

Evaluative studies of the interface software developed in this project were conducted employing user questionnaires, transaction log analysis and focused group interviews. A statistical analysis of user questionnaires and search results has been conducted looking for relationships between demographic variables, search satisfaction and search results. The project examined interface design features, particularly measures to facilitate search strategy formulation and software mediated strategy modification.

The questionnaires administered to users gathered background information, including age, gender, university status and major field of study; previous library and online catalog use; previous computer experience; and previous searching experience. The questionnaire also asked users to report satisfaction with search results, percentage of relevant citations, ease of use, problems associated with online searching, comparison with printed secondary tools and ability to use the search software. A site monitor form was also filled out for each search session, reporting the number of times assistance was requested and the type of assistance, user's ability to use the keyboard and search system and enthusiasm for searching. A statistical analysis of some 500 user questionnaires covering over 1,000 searches has been conducted, compiling frequencies and setting up cross-tabulations of the variables using a Chi-Square analysis. Overall, 96 percent of the users rated the system easy to use, 89 percent were satisfied with search results, and 62 percent of the searchers indicated that one-half or more of retrieved references were relevant.

The transaction log analysis and focused group interviews have provided insight into end-user searching behavior.

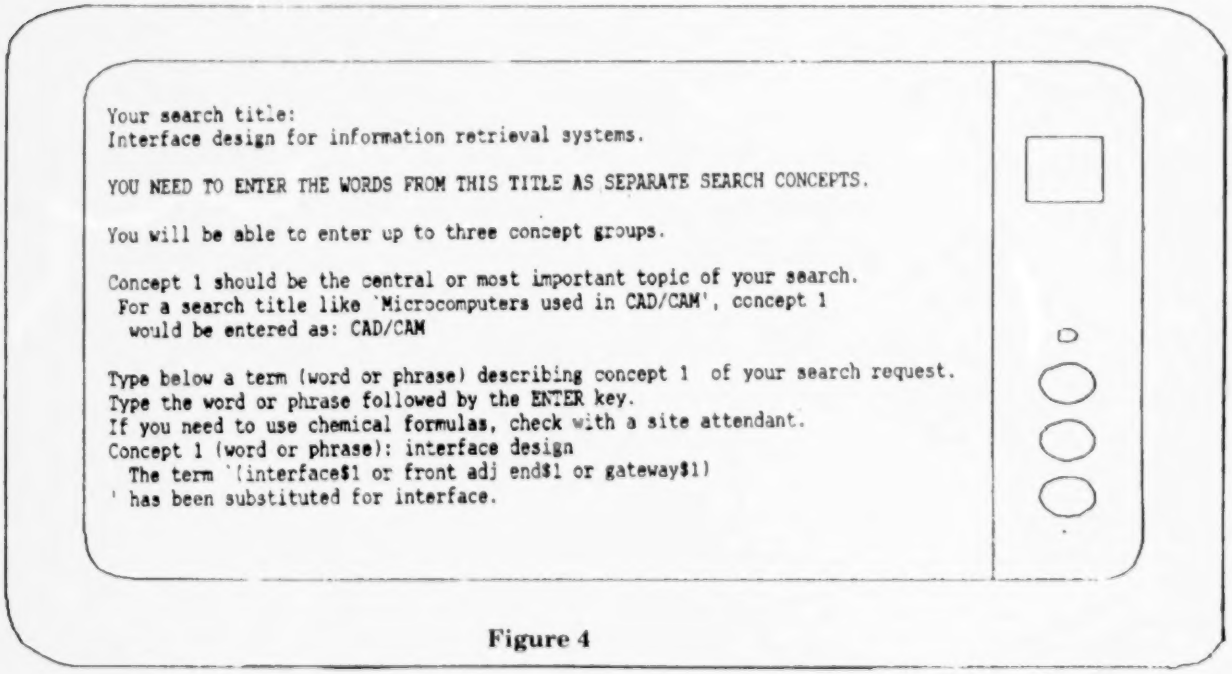

Your search title:
Interface design for information retrieval systems.
YOU NEED TO ENTER THE WORDS FROM THIS TITLE AS SEPARATE SEARCH CONCEPTS.
You will be able to enter up to three concept groups.
Concept 1 should be the central or most important topic of your search.
For a search title like 'Microcomputers used in CAD/CAM', concept 1
would be entered as: CAD/CAM
Type below a term (word or phrase) describing concept 1 of your search request.
Type the word or phrase followed by the ENTER key.
If you need to use chemical formulas, check with a site attendant.
Concept 1 (word or phrase): interface design
The term '(interface$1 or front adj end$1 or gateway$1)
' has been substituted for interface.

Figure 4

The transaction log analysis and focused group interviews have also provided a great deal of insight into end-user searching behavior. Among the findings:

1. A truncation algorithm is necessary. Users tend not to think about entering plural forms.
2. The "NOT" operator is needed in a small but significant number of searches.
3. Users are unaware of basic set-theoretic concepts; they often enter specific terms that are subsumed within already-entered broader terms (e.g., "earth metals" OR "alkaline earth metals").
4. Users enter very specific multiword phrases as search terms. It is important for the interface to employ less restrictive proximity operators (WITH/NEAR as opposed to ADJ).
5. Users have problems with selection of entry terms or search vocabulary; any online vocabulary assistance is useful.
6. The software manipulation of resultant search sets has proven useful.
7. Many users are interested in retrieving a relatively small number of references; however, limiting retrieval to 25 references is a problem for some users.
8. Online modification of search strategy is needed, but is not trivial to set up and perform in some searches.

CONCLUSION

The results of the end-user and online catalog use studies have important ramifications in the design and development of effective user interfaces for bibliographic retrieval systems. The interface plays a particularly critical role in the searching of bibliographic retrieval systems (such as periodical index databases) that employ sophisticated information retrieval techniques and contain records with subject-rich fields, including abstracts, descriptors and identifiers.

Clearly, advances in interface design can be applied to online catalogs, online periodical citation databases and optical disk search systems. Providing assistance to users in search strategy formulation and modification appears to be the area of greatest need in interface design.

Overall, this project has demonstrated that end-users are willing to use a menu-driven interface containing instructional material and can be assisted in the performance of effective online searches of bibliographic databases. The project has identified several areas of interface design that need further development.

REFERENCES

[1] Arms, Caroline, "The Technological Context" in Campus Strategies for Libraries and Electronic Information, ed. by Caroline Arms, Digital Press, 1989. pp. 33-34.

[2] Hildreth, Charles R., Intelligent Interfaces and Retrieval Methods for Subject Searching in Bibliographic Retrieval Systems, Library of Congress, 1989.

[3] Mischo, William H., and Lee, Jounghyoun, "End-User Searching of Bibliographic Databases," Annual Review of Information Science and Technology, v22, 1987. pp. 227-63.

[4] Markey, Karen, "Thus Spake the OPAC User," Information Technology and Libraries, v2, December 1983. pp. 381-387.

[5] Hildreth, Charles R., "Beyond Boolean: Designing the Next Generation of Online Catalogs," Library Trends, v35, Spring 1987. pp. 647-658.

[6] Hildreth, Charles R., "Online Public Access Catalogs," Annual Review of Information Science and Technology, v20 1985. pp. 233-285.

[7] Matthews, Joseph R., and Lawrence, Gary S., "Further Analysis of the CLR Online Catalog Project," Information Technology and Libraries, v3, December 1984. pp. 354-376.

[8] Mischo, William H., and Moore, Amy F., "Enhanced Access to Periodical Literature Within an Online Catalog Environment" in Online Catalogue Research: Developments and Directions, ed. by Charles R. Hildreth, London, Library Association, 1989. pp. 107-126.

NEW CHALLENGES IN HELPING STUDENTS UNCOVER INFORMATION: THE LEARNING DISABLED STUDENT

By Marian Traina Power, Director,
Carol Rundlett, Assistant Director, and
Myra David, Graduate Student,
Academic Resource Center, C.W. Post Campus
Long Island University, Brookville, NY

In 1976 Marian Power described the "new student" on the postsecondary scene.[1] This new student had "external" disadvantages stemming from poverty, neglect and/or cultural isolation. Some 13 years later, the "new student" for the 1990s appears to be one with an internal rather than external disadvantage. These students have been classified as "learning disabled" (LD).

It is imperative that college librarians become aware of the needs of this new population.

Current statistics regarding LD students are limited since many students do not identify themselves as being LD. According to the American Freshman Norms for 1986, over 14 percent of all freshmen with disabilities reported they were LD.[2] Increasing numbers of students with learning disabilities are seeking postsecondary education.[3,4] For educational as well as legal reasons, it is imperative that college librarians become aware of the needs of this new population. Librarians must be accessible and prepared to help LD students use every service the library offers.

This article is intended to familiarize librarians with the characteristics of LD college students. Librarians can then develop a sensitivity to this new population requesting their services.

Improved identification and greater support services for LD students at the high school level will result in greater numbers of individuals with learning disabilities as potential consumers of postsecondary services in the 1990s. But there are legal reasons for this new population as well.

Section 504 of the Federal Rehabilitative Act of 1973 (Public Law 94-142) is the civil rights law protecting the rights of individuals with disabilities. It has widespread application as learning disabled students apply for admission to colleges and universities in the United States. Congress assigned to the U.S. Department of Health, Education and Welfare the responsibility to see that disabled individuals will be able to participate in programs and activities receiving Federal financial assistance and student aid (a definition that includes nearly all colleges and universities). Since enactment of Section 504, the composition of American colleges and universities continues to change as more and more students are revealing their learning disabilities and demanding the appropriate support services to which they are legally entitled.

The Department of Health, Education and Welfare in 1977 defined a learning disability as a disorder in one or more of the basic psychological processes involved in understanding or in using language, spoken or written, which manifests itself in an imperfect ability to listen, think, speak, read, write, spell or do mathematical calculations. The Federal regulations pertaining to serving the learning disabled student indicate that:

> Those with specific learning disabilities may demonstrate their handicap through a variety of symptoms such as hyperactivity, distractibility, attention problems, concept association problems....The end result of the effect of these symptoms is a severe discrepancy between achievement and ability.[5]

While Public Law 94-142 remains the most widely used definition of learning disabilities,[6] there are a number of other definitions. The common thrust of these is that learning disabilities are not a single disability but made up of several subtypes of handicapping conditions.[7]

Neither a casual look nor an intense look at an individual reveals specific learning disabilities. Librarians should not and can not have the responsibility of recognizing the learning disabled student.

The United States Department of Education reported in 1984 that the number of programs for learning disabled students had doubled within the last 10 years and that more than 10 percent of all pupils served in special education programs were classified as being learning disabled. According to data on vari-

ous states in this report, learning disabilities ranged from 4 to 10 percent of the special education population.[8] Today, the number of individuals identified as learning disabled and the number of new programs for them continues to increase. These postsecondary programs generally fall into three categories:

- a prep program (a precollege year, semester or summer focusing on learning styles, study skills and organization)
- an LD program (with separate admissions and at least some courses separate from the nondisabled population)
- a regular campus program (in which the students are mainstreamed into regular college classes but have special support services offered to them)

The library's participation in these programs is meager.[9] However, Donna Corlett found that library skills could be used as a predictor of college success,[10] and Allard, Dodd and Peralez cite faculty awareness and support of LD students and their needs first in their solutions to the problem of keeping LD students in college.[11] Perhaps with more awareness and support for LD students on the part of the staff of college libraries, the success rate of LD college students can be increased.

When an individual approaches a librarian and indicates that she or he is learning disabled, what information is that person communicating?

When an individual approaches a librarian and indicates that she or he is learning disabled, what information is that person communicating? In general, learning disabled students have average or above average intelligence and many are able to verbally organize their thoughts but are unable to achieve their goal in the library. They become confused because they have difficulty selecting the main idea of their topic or are unable to categorize their ideas. Many learning disabled students have difficulty sorting relevant from irrelevant information.

Academically, the most common problem associated with learning disabled students is difficulty reading. However, this is not the only disability librarians may encounter.[12] Dyscalculia (difficulty with mathematics) may interfere with a student's ability to use the Library of Congress classification system to find a book on the shelf or to find a correct page in a reference book.

Learning disabled students may have difficulty remembering the correct formation of letters and have other spatial problems.[13] In the library, this may mean an inability to copy down information needed to find reference material or to find information from reference material.

Learning disabled students may also have long-term memory problems that require the librarian to reinstruct the learning disabled student in the use of equipment or materials several times. It is not that the student is not listening the first time — rather, it is that LD people often have difficulty moving information from short-term to long-term memory.

Spatial disabilities may mean a student needs much more than a walking tour of the library during freshman orientation. She or he may need an audiotape map or a textured map depending on his/her learning style. The student may also have tremendous difficulty using a microfilm reader, which requires strong eye/hand coordination as well as left/right, up/down orientation.

The librarian working with an LD college student must be aware that there are differences in the basic processing abilities for these young adults. The librarian needs to help the student focus and maintain attention on the specific task long enough to complete the task. Forster and Doyle found that students with learning disabilities have never learned to listen effectively or they have developed poor listening habits.[14]

A short attention span may mean a learning disabled student cannot learn the card catalog and the stacks on the same day. She or he may not be able to concentrate on library lectures or may be distracted by common library sounds and noises.

The student may struggle for hours with the card catalog, copy machine or encyclopedia before giving up and going back to the dorm having achieved nothing but another failure.

Finally, a learning disabled student may have had

so many bad experiences with teachers, librarians and peers who mistook LD problems as symptoms of laziness or stupidity and refused help that the student may be afraid to come up to a librarian at all. The student may struggle for hours with the card catalog, copy machine or encyclopedia before giving up and going back to the dorm having achieved nothing but another failure.

Librarians can help LD college students by:

1. Becoming aware of this new population and its needs. At Long Island University, C.W. Post Campus, we recommend sharing this information through publication and staff development. Minner and Prater recommend staff development workshops to combat the negative influence which they found of labelling students on the attitudes of college faculty and staff.[15]
2. Being available to LD students on an as-needed or appointment basis. Many LD students are learning about time management techniques and could benefit from a specific appointment. Two good strategies are making contact with whatever program the school offers for LD students and letting students know there is a sincere desire to help them.
3. Considering the needs of LD students when designing or selecting an online catalog. The catalog might produce a range of text magnifications; graphics and color could be used to enhance the display. Both large print and perhaps adjustable scrolling of the text (horizontally or vertically) would be beneficial. Other text attributes such as bold print, underlining and arrows could be used to aid LD students.
4. Considering the needs of LD students when instructing in the use of a card catalog or online catalog. Spatial, sequencing, reading and math problems all make the card catalog a seemingly insurmountable challenge for some learning disabled students. Students should not become so discouraged and frustrated with the catalog that they avoid going to the library.
5. Marking stack shelves clearly and providing assistance in the stacks for those with spatial and other pertinent disabilities.
6. Providing extra assistance with microfilm readers and other equipment.
7. Investigating to what extent librarians can effectively serve LD students within the structures of the library. Research is needed to evaluate the effect of different library settings on the elementary, secondary and postsecondary levels.
8. Applying for grants and funding to address ways to evaluate the progress of students with learning disabilities in a variety of educational settings including the library.
9. Finally, but most importantly, conveying a feeling of respect for those learning disabled students struggling to get an education and showing them that librarians are there as a resource and to help them achieve success in college.

Librarians need to implement well planned, well structured programs for college students with learning disabilities. The nine steps described in this paper are key components for doing so. Only when the library is a place of success rather than frustration and failure for the LD college student, can college librarians feel they are serving this new and growing population.

REFERENCES

1 Power, Marian, "New Students Mean Change for Colleges," *Journal of Reading*, v20, n3, December 1976. pp. 237-240.

2 Heath Resource Center, "Learning Disabled Adults in Postsecondary Education," 1987 Edition. pp. 1-8.

3 Longo, Judith, "The Learning Disabled: Challenge to Postsecondary Institutions," *Journal of Learning Disabilities*, v11, n3, Spring 1988. pp. 10-14.

4 Minner, Sam, and Prater, Greg, "College Teachers' Expectations of LD Students," *Academic Therapy*, v20, n2, November 1984. pp. 225-29.

5 *Federal Register*, v42, n250, Thursday, December 29, 1977. pp. 65082-5.

6 Lazarus, Belinda D, "Serving Learning Disabled Students in Postsecondary Settings," *Journal of Developmental Education*, v12, n3, January 1989. pp. 2-6.

7 McKinney, J., "The Search for Subtypes of Specific Learning Disabilities," *Journal of Learning Disabilities*, v7, n3, Spring 1984. pp. 43-50.

8 U.S. Department of Education, *Sixth Annual Report to Congress*, Washington, DC, Government Printing Office, 1984. pp. 28-30.

9 Adler, Rebecca, "Unlocking the Library for the Learning Disabled Student," *Issues in College Learning Centers*, v6, Winter 1988. pp. 36-42.

10 Corlett, Donna, "Library Skills, Study Habits and Attitudes and Sex as Related to Academic Achievement," *Educational and Psychological Measurement*, v34, March 1974. pp. 967-69.

11 Allard, W. Gregory, Dodd, John M., and Peralez, Esther, "Keeping LD Students in College," *Academic Therapy*, v22, n4, March 1987, pp. 359-65.

12 Becker, Sidney G., "The Library/Media Center for the Secondary Student with Specific Learning Disabilities — Hazard or Haven?" in *The Mainstreamed Library: Issues, Ideas, Innovations*, Chicago, American Library Association, 1982, pp. 46-9.

13 Becker, *Op. Cit.*

14 Forster, Patricia, and Doyle, Beverly, "Teaching Listening Skills to Students with Attention Deficit Disorders," *Teaching Exceptional Children*, Winter 1989, pp. 20-23.

15 Minner and Prater, *Op. Cit.*

THREE-PART PLAN TO IMPROVE LIBRARY SERVICES APPROVED

The Board of Regents in January 1989 endorsed five standards to improve public library services in New York State and asked Education Commissioner Thomas Sobol to do further work before proposing revisions to current library regulations.

According to Regent Laura B. Chodos, Chair of the Regents Cultural Education Committee, "the quality of library service today depends in large part upon where you live. To change this, the Regents will update regulations and seek aid to assure equity of access to library services."

"Before the Regents take action on any regulations," she added, "hearings will be held." A three-part strategy adopted by the Regents based on the recommendations of Commissioner Sobol includes the following elements:

- Five recommendations, requiring accountability but not requiring significant out-of-pocket expense, were endorsed by the Regents, expecting that they can be put into Commissioner's Regulations late in 1989. These would require a library to have bylaws, a plan of service, written policies and an annual budget, and to make available printed information about hours, borrowing policies and the library telephone number and address.
- Four other recommendations, which would require significant expenditures by some libraries, will be the subject of consultation between the Commissioner and concerned parties to develop funding strategies. Under these recommendations, libraries would provide services to the disabled, connect with automated services, provide rest rooms, photocopiers and telephones. The four standards would require a one time cost of $2 million to bring all 737 libraries in the State up to standard and an ongoing annual cost of $560,000. The Regents asked staff to consult with relevant parties on funding strategies to enable libraries to meet these standards, with the "understanding that the process does not bind the Regents at this time to take any specific action" on them.
- The Regents did not take action on an earlier recommendation concerning certification of librarians for small libraries. The Regents asked Education Department staff to work with rural library staff and trustees, public library system directors, library schools, the Regents College Degree program and providers of distance learning to develop appropriate training and certification procedures for small libraries. According to Regent Chodos, the Regents expect staff to do this work within three years.

Current minimum standards for public libraries have not been revised since 1950. The Regents have been responsible for setting minimum standards of service for public libraries as educational institutions since the 1890s.

CLARKSON MAKES NYSERNet AN INTEGRAL PART OF LIBRARY SERVICE

Network services — A luxury? — An important part of the resource mix? — A necessity? For Clarkson University's Educational Resources Center, NYSERNet is becoming an integral part of the resource base available to library users, and it sees NYSERNet as vital to its ability to provide adequate library services in the next decade.

There are two reasons for Clarkson's special interest in electronic library services. A decade ago, its board of trustees became convinced that information would be created, manipulated, stored and transmitted electronically to even greater extent into the next century. Therefore, the trustees designed Clarkson's new library to facilitate electronic services, and they combined library, computing and media services into one administrative unit, the Educational Resources Center. The library has benefited from both this administrative interest in promoting electronic services and from the computing and networking expertise available from colleagues.

Clarkson's relatively remote location, three hours from the nearest metropolitan center and from any library with a substantial science and engineering collection, has also contributed to their interest in electronic information sources. The rapid inflation in journal prices over the past 10 years has increased their determination to take full advantage of the opportunities that NYSERNet offers.

The Colorado Alliance for Research Libraries offered a unique opportunity for Clarkson to reach an important and growing resource base. CARL's position on the Colorado Supernet allowed them to use their NYSERNet connection for direct faculty/student desktop access as well as for access through library workstations. Currently, we connect to the library catalogs of most Colorado libraries, to the Academic American Encyclopedia and to CARL's unique *UNCOVER* database. Starting in the summer of 1988, *UNCOVER* includes the contents pages of all magazines and journals received by Colorado libraries — about 10,000 titles in all. Citations to the journal articles can be retrieved individually by author or title keyword search, or can be reassembled as the table of contents for any given journal issue.

Under CARL's dynamic leadership, Clarkson expects to see a steady development of new services and applications, including more databases, bulletin boards, consultant files, full-text journal articles and electronic transmission of facsimile images over the network. CARL also offers customized services. An institution, for example, could arrange a customized *UNCOVER* database comprising only the journals to which they themselves subscribe. More importantly, other institutions are also beginning to offer information services available over network connections.

Public access workstations connected to NYSERNet now compete with CD-ROM workstations, online database access and, of course, traditional printed works in Clarkson's library reference area. The Educational Resources Center is currently working on networking arrangements that would allow each library workstation to access a wide array of electronic services, including NYSERNet services. Also important is enhancing access to library services for individual researchers in their offices. Clarkson has currently set up shell scripts that allow any individual on their campus network to automatically make the CARL connection simply by typing CARL at the system prompt. Managing this type of information access is a key concept in Clarkson's vision of the future of library and information services.

This article by J. Natalia Stahl, Assistant Director of the Educational Resources Center at Clarkson University, originally appeared in the July-August 1989 issue of NYSERNet News *and is reprinted here with permission.*

The New York State Library

Established in 1818, the New York State Library is the largest state library in the nation and the only state library to qualify for membership in the Association of Research Libraries. It is a leader in the use of technology in information services, with an online public catalog, an electronic reference station, access to more than 600 databases, including RLIN and OCLC, and modern telefacsimile and photoduplication services.

The State Library's research collection of more than five million items includes major holdings in law, medicine, the social sciences, education, American and New York State history and culture, the pure sciences, and technology. Its information resources include manuscripts, government documents, newspapers, microforms, patents, and bibliographic aids, including abstracting, indexing, and loose-leaf services.

The State Library provides reference and research services to support the work of New York State government. It serves as a resource and referral center in the New York State Interlibrary Loan Network, and operates a regional library for the blind and visually handicapped. The State Library is also the headquarters of the New Netherland Project.

The New York State Library is a modern research facility occupying six levels of the Cultural Education Center in the Nelson A. Rockefeller Empire State Plaza. The Library is accessible to the handicapped.

NEW YORK STATE LIBRARY

Manuscripts and Special Collections

Manuscripts and Special Collections was established as a unit within the New York State Library in 1881. The function of this section is to acquire, arrange, describe, catalog, and preserve manuscripts, rare books, maps, and prints, broadsides, microfilm, sheet music, and ephemera which illustrate the history of New York State from the 17th century to the present and to provide reference and research assistance to State government, the scholarly community and the general public. The unit is located on the 11th floor of the Cultural Education Center.

RULES AND REGULATIONS

Users of the research facilities of the Manuscripts and Special Collections must register daily when using the collections and must follow the research room regulations provided at registration. Various catalogs, automated and manual, provide detailed access to the collections.

Librarians are responsible for the determination of which materials may be available for use. Because of donor or legal restrictions or fragile physical condition of the material, access cannot be assured to all holdings.

Microfilm readers and printers are available and the use of computers by researchers is allowed. Photoduplication work is done by staff only and the reproduction of any material may be done only after consultation with staff. The Library does not assume any responsibility for literary or publication rights of material in its collection. Researchers should retain such permission from the owners of such rights.

Because of the valuable and unique nature of the collection nearly all materials are noncirculating and must be used in the research room.

Manuscripts

The New York State Library has approximately 19,000 single manuscript accessions and over 1,000 manuscript collections ranging in size from one linear foot to over 2,000 cubic feet. These manuscripts date from 1630 and relate primarily, though not exclusively, to the history of the State and its people.

MAJOR TYPES OF MATERIAL

- correspondence
- diaries and journals
- business and legal records
- land records, including leases, deeds and surveys
- essays and literary papers
- sketch books
- scrapbooks
- church and genealogical records
- estate papers, including wills and inventories
- photographic and printed materials as part of manuscripts collections
- institutional and organizational records
- maps as part of manuscripts collections

HISTORICAL COLLECTIONS

Materials support research in many aspects of New York State and United States history including:

- military history (John Ellis Wool and G.K. Warren papers)
- economic history (Burden Iron Company and Delaware and Hudson Railway records)
- political history (Edwin D. Morgan, Horatio Seymour, Alfred E. Smith papers)
- history of science (James Hall and Amos Eaton papers)
- art history (Thomas Cole and John F. Kensett papers)
- family-social history (Historic Cherry Hill, Tibbits Family, Lansing family papers)
- organizational history (New York Library Association and New York Child Labor Board records)
- institutional history (Parsons Center records)
- religious history (church records of several Protestant denominations)
- environmental history (Adirondack Council and Nature Conservancy papers)
- microfilmed manuscripts and archives found in other libraries (Gerrit Smith papers, George Washington papers, and Shaker manuscripts)

GOVERNMENT RECORDS

The Library does not collect public or official government records; these are held by the State Archives and Records Administration.

TREASURES

There are several very significant documents of national historic importance including Lincoln's Preliminary Emancipation Proclamation and the first draft of Washington's Farewell Address.

Cartographic Material

The use of cartographic material is a vital part of research by land planners, environmentalists, historic preservationists, genealogists, surveyors, lawyers, geologists, hikers, historians and others. The State Library's Manuscripts and Special Collections Section staff, through the use of the Library's extensive collection of maps and atlases, seeks to aid these researchers in solving their research problems.

The Library collects a variety of cartographic materials by purchase, gift and through its status as a depository for State and Federal documents. The geographic emphasis of the collection is New York State and the northeastern United States. In addition, the Library has maps of the United States, the world and specific areas and nations of the world in a variety of scales.

MAPS

Some of the important map series available for research include:

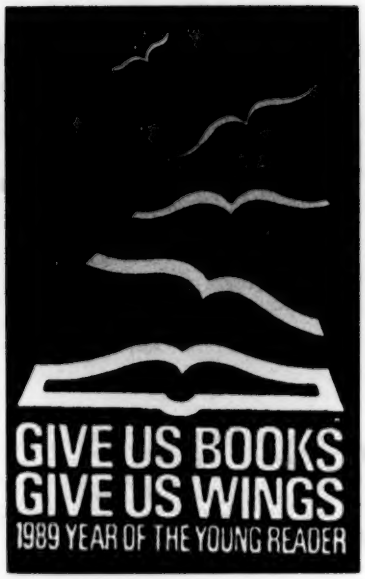
- Comprehensive collection of maps of New York State and its cities, counties, and towns from the mid-17th century to the present

- Large numbers of maps of the New England states and the United States from the mid-18th century to the present
- United States Geological Survey topographical maps of the United States, current and historical, at various scales
- United States Geological Survey geological maps for various subjects including oil and gas investigations and hydrologic reports
- National Ocean Survey, Great Lakes survey charts
- Plastic relief maps of the United States
- United States Soil Conservation Service's soil maps
- United States. Defense Mapping Agency. Maps of the world in various scales
- New York State Museum's Map and Chart Series
- James Hall Papers' manuscript geological maps of New York in the 19th century
- Historic Cherry Hill's 19th century maps of eastern New York State and the Capital District region
- Van Rensselaer Manor Papers' late 18th and early 19th century survey maps of Albany and Rensselaer counties

ATLASES

The atlas collection numbers over 1,000 atlases paralleling collecting themes of the map collection. Included in this collection are such notable historical atlases as:

- comprehensive set of New York State county atlases
- J.F.W. DesBarres' *Atlantic Neptune* (1777-1781)
- Thomas Jeffreys' *American Atlas* (1776) and *West India Atlas* (1775)
- William Faden's *North American Atlas* (1777)
- Christopher Colles' *Survey of the Roads of the United States* (1789)
- Pieter Goos' *L'Atlas de la Mer* (1667)
- David H. Burr's *Atlas of the State of New York* (1829, 1839)
- Sanborn insurance atlases for several New York State cities and the New York State insurance atlas collection at the Library of Congress on microfilm



The New York State Library is a primary repository for publications pertaining to the history and development of New York State and would appreciate notification from librarians, authors and historical societies of publications available by purchase or gift. Correspondence should be directed to: Humanities Reference Service, New York State Library, Cultural Education Center, Albany, NY 12230.

All articles in *The Bookmark* are published on the authority of the authors. They do not necessarily reflect the opinion of the editor or the policy of the State Education Department.

Reproduction of articles from *The Bookmark* for educational use, with appropriate attribution of source, is permitted, except that permission to reproduce copyrighted articles must be obtained from the copyright holder identified on the first page of the article.

Credits: In addition to editors and others credited elsewhere in this issue, the following persons in the New York State Library helped make this issue possible: Christopher Burns, Production Manager; Andrew Mace, Editorial Assistant; and Christine Hodor, Manuscripts Control.

New York State Education Department
New York State Library
Cultural Education Center
Albany, New York 12230

NONPROFIT ORG.
U. S. POSTAGE
PAID
Albany, N.Y.
Permit No. 293

89-7988
CAD # 90-088

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991